

TARTU ÜLIKOOL
LOODUS- JA TEHNOLOOGIATEADUSKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
BIOINFORMAATIKA ÕPPETOOL

Kairi Raime

**Taksoni-spetsiifiliste praimerite disaini
metoodika arendus ning selle rakendamine
allergeensete taimeliikide tuvastamisel**

Magistritöö

Juhendajad prof. Mairo Remm
PhD Triinu Kõressaar

TARTU 2015

Sisukord

Sisukord.....	2
Kasutatud lühendid.....	4
Sissejuhatus	5
1. Kirjanduse ülevaade	7
1.1. Polümeraasi ahelreaktsioon ja selle efektiivsus.....	7
1.2. Amplifikatsiooni efektiivsust mõjutavad tegurid	8
1.2.1. Praimeri ja DNA järjestuse vahelise hübridisatsiooni tugevus	8
1.2.2. DNA polümeraasi seondumise ja praimerite pikendamise efektiivsus	10
1.3. Taksoni-spetsiifiliste testide väljatöötamine.....	16
1.4. Olemasolev tarkvara praimerite spetsiifilisuse kontrollimiseks	18
1.4.1. Tarkvara praimerite mittespetsiifiliste seondumiste ennustamiseks	19
1.4.2. Tarkvara mittespetsiifilise amplifikatsiooni ennustamiseks.....	20
2. Eksperimentaalne osa	23
2.1. Töö eesmärgid.....	23
2.2. Materjal ja metoodika	24
2.2.1. Mittepaarduvate nukleotiididega praimerid	24
2.2.2. DNA eraldus.....	28
2.2.3. Polümeraasi ahelreaktsiooni katsed	28
2.2.4. Arvutitarkvara arendus taksoni-spetsiifiliste praimerite disainiks	29
2.2.5. Taksoni-spetsiifiliste praimerite disaini metoodika rakendamine diagnostilise testi väljatöötamisel	30
2.3. Tulemused.....	32
2.3.1. Skoor praimerite pikendamise ennustamiseks	32
2.3.2. PCR katsete tulemused üksikuid mittepaarduvaid nukleotide sisaldavate praimeritega.....	34
2.3.3. Seos amplifikatsiooni saagise ja praimeripaari NAS väärtuse vahel	35
2.3.4. Taksoni-spetsiifiliste PCR praimerite disaini metoodika ja tarkvara.....	39
2.3.5. Väljatöötud tarkvara ja metoodika abil disainitud spetsiifilised praimerid pekanipähkli tuvastamiseks.....	43
2.4. Arutelu	46
Kokkuvõte	51
Summary.....	52
Tänuavaldused	54

Kirjanduse loetelu.....	55
Kasutatud veebiaadressid	60
Lisad	61
Lisa 1.....	61
Lisa 2.....	63
Lisa 3.....	66
Lisa 4.....	69
Lisa 5.....	73

Kasutatud lühendid

bp – aluspaar (**b**ase **p**air)

NAS – praimerite seondumise ja pikendamise efektiivsust iseloomustav skoor (*No Amplification Score*)

NCBI – USA Riiklik Biotehnoloogilise Informatsiooni Keskus (*National Center for Biotechnology Information*)

PCR – polümeraasi ahelreaktsioon (*p*olymerase *c*hain *r*eaction)

T_m – sulamistemperatuur (*melting temperature*)

Sissejuhatus

Vaatamata sekveneerimismeetodite üha laiemale levikule on polümeraasi ahelreaktsioon (PCR) tänu oma odavusele, kiirusele, lihtsale tööpõhimõttele, spetsiifilisusele ja tundlikkusele eelistatud meetodina kasutusel paljude erinevate diagnostiliste testide puhul (näiteks patogeensete/allergeensete/mürgiste/ohustatud taksonite tuvastamine toiduainetest, keskkonnaproovidest). PCR-il põhineva taksoni-spetsiifilise diagnostilise testi puhul on eesmärk konkreetse taksoni (nt. alamliik, liik, perekond, sugukond) DNA tuvastamine sarnaste järjestustega sihtmärk- ja mittesihtmärk-taksonite DNA-de segust, milles sihtmärk-DNA osakaal võib olla väga väike.

Kaks kõige olulisemat omadust PCR-il põhineva taksoni-spetsiifilise testi puhul on selle spetsiifilisus (ainult sihtmärk-taksoni DNA tuvastamine ja mitte ühegi samas DNA proovis esineva mittesihtmärk-taksoni DNA tuvastamine) ja tundlikkus (väga madala kontsentratsiooniga sihtmärk-taksoni DNA tuvastamine), mille aluseks on kasutatavad PCR praimerid ja nende sihtmärk- ning mittesihtmärk-DNA järjestusega seondumise ja pikendamise efektiivsus.

Taksoni-spetsiifilistes testides kasutatavate praimerite disain hõlmab endas lisaks oligonukleotiidsete järjestuste genereerimisele ka eelnevat sihtmärk- ja mittesihtmärk-taksonite määramist, disainiks kasutatava genoomipiirkonna valikut, järjestuste hankimist andmebaasidest ja nende võrdlemist ning pärast praimerite genereerimist ka *in silico* praimerite sobivuse (spetsiifilisuse, universaalsuse) kontrolli ja konkreetse taksoni-spetsiifilise testi jaoks kõige sobivama(te) praimeripaari(de) väljavalimist.

Praimerite disainiks taksoni-spetsiifilise testi jaoks on vaja piisavalt palju nii sihtmärk-taksoni genoomseid järjestusi (et teada sihtmärk-taksoni sisest järjestuste varieeruvust) kui ka mittesihtmärk-taksonite genoomseid järjestusi (et leida sihtmärk-taksoni puhul unikaalseid kohti praimerite disainiks). Kuigi genoomijärjestuste andmebaasides on üha rohkem täisgenoomijärjestusi, on veel väga palju organismirühmi (nt taimed), mille puhul sekveneeritud genoomijärjestusi on veel liiga vähe, et neid taksoni-spetsiifiliste praimerite disainil kasutada. Kuna disainiks kasutatavate geenipiirkondade ja järjestuste hulk andmebaasides on väga piiratud, on vähe infot taksoni-sisese järjestuste varieeruvuse ning sihtmärk- ja mittesihtmärk-taksonite järjestuste vahelise varieeruvuse kohta. Kui taksoni-spetsiifilise testi eesmärk on tuvastada allergeenseid või toksilisi aineid sisaldavate organismide taksoneid toiduainetest (nt. allergeensed taimed, loomad) ja oluline on testi tundlikkus, on oluline vältida sihtmärk-taksonite järjestuste väikestki varieerumist praimerite seondumissaidis.

Samas võib mõnede organismirühmade (nt bakterite või mõnede majanduslikult oluliste taime- või loomarühmade) puhul olla probleem, et sihtmärk-taksoni järjestusi on andmebaasis küll piisav hulk, kuid järjestused varieeruvad sihtmärk-taksoni siseselt liiga palju ja samal ajal liiga vähe võrreldes fülogeneetiliselt lähedaste mittesihtmärk-taksoni järjestustega. Sellises olukorras võib taksoni-spetsiifiliste praimerite disainiks sobivate kohtade leidmine olla väga keeruline.

Praimerite spetsiifilisus sõltub olulisel määral sellest, kuivõrd efektiivselt praimer mittesihtmärkjärjestusega seondub ning kuivõrd efektiivselt moodustunud dupleksit pikendatakse, mis omakorda sõltub mittepaarduvate aluspaaride arvust, paiknemisest ja tüübist praimer ja mittesihtmärkjärjestuse vahelises dupleksis.

Antud töö kirjanduse osas antakse kõigepealt ülevaade PCR amplifikatsiooni efektiivsust mõjutavatest ja selle ennustamiseks kasutatavatest praimeritega seotud teguritest (praimerite ja DNA järjestuse vahelise hübridisatsiooni tugevus ja praimerite pikendamise efektiivsus). Seejärel antakse ülevaade taksoni-spetsiifiliste testide disaini eripäradest, olulistest aspektidest ja väljakutsetest praimerite disainil diagnostilise testi jaoks ning olemasolevatest arvutiprogrammidest taksoni-spetsiifiliste praimerite disainiks, mittespetsiifiliste seondumiste või amplifikatsiooni ennustamiseks ning olemasoleva tarkvara puudustest.

Antud töö eksperimentaalse osa peamiseks eesmärgideks on töötada välja mittepaardunud aluspaaride arvul ja positsioonil baseeruv praimerite seondumise ja pikendamise efektiivsust kirjeldav skoor, kontrollida praimeripaari skoori väärtuse seost amplifikatsiooni saagisega, rakendada väljatöötatud skoori praimeripaari spetsiifilisuse hindamiseks ning luua tarkvara ja meetodika taksoni-spetsiifilise testi jaoks praimerite disainiks ja valimiseks Unix/Linux keskkonnas. Väljatöötatud meetodikat ja tarkvara katsetati liigispetsiifiliste PCR praimerite disainil allergeense taimeliigi pekani-hikkoripuu (mille vili on pekanipähkel ehk pekaani pähkel) tuvastamiseks Allergofood projekti raames.

Märksõnad:

taksoni-spetsiifilised praimerid (*taxon-specific primers*), skoor (*score*), spetsiifilisuse kontrollimine (*specificity checking*), tarkvara (*software, program*).

1. Kirjanduse ülevaade

1.1. Polümeraasi ahelreaktsioon ja selle efektiivsus

Polümeraasi ahelreaktsioon (PCR) on molekulaarbioloogias üks kõige laiemalt kasutusel olev meetod, mis oma tundlikkuse tõttu võimaldab paljundada ka väga väikeses algkoguses geneetilist materjali, et tuvastada erinevate organismide (nt patogeenide, allergeensete või toksiliste liikide, viiruste) DNA olemasolu, geneetilisi variatsioone (mutatsioone) inimese geenides või saada suurem kogus DNA materjali muudeks rakendusteks.

Ideaalsetes tingimustes kahekordistatakse iga järgneva PCR tsükli käigus eelneva tsükli DNA hulk. N tsükli möödudes on esialgse kaheaahelalise sihtmärk-DNA koopia arv suurenenud 2^n korda. Et DNA proovi kontsentratsiooni piisavalt suurendada, korratakse PCR tsükli tavaliselt 20 - 35 korda. Seega 35 tsükli järel on esialgset DNA-d sünteesitud 2^{35} ehk 34 miljardit koopiat. (Whitney *et al.*, 2004; Chuang *et al.*, 2013). Produkti hulk igas PCR tsükli on ideaalsel juhul $N = N_0 \times 2^n$, kus N on amplifitseeritud molekulide arv, N_0 on esialgne molekulide arv, n on amplifikatsiooni tsüklite arv (Cha ja Thilly, 1995; Wiesner, 1992; Higuchi *et al.*, 1993).

Tegelikkuses ei tarvitse amplifikatsiooni efektiivsus olla 100% ning võib kõikuda vahemikus 0-st 1-ni (1 tähendab 100% efektiivsust, mille puhul molekulide arv tsükli käigus kahekordistub), seega on tegelik PCR lõpp-produkti hulka kirjeldav valem $N = N_0 \times (1 + E)^n$, milles lisamuutuja E on amplifikatsiooni efektiivsus (Arezi *et al.*, 2003). Sõltuvalt kasutatavast DNA polümeraasist, reaktsioonitingimustest ja sihtmärk-DNA järjestusest (nt G ja C nukleotiidide osakaalust järjestuses) võib 10^9 kordse amplifikatsiooni saavutamiseks vajalik minimaalne PCR tsüklite arv olla seetõttu ka tavapärasest 30 tsüklist oluliselt suurem. Mida suurem on aga PCR tsüklite arv, seda suurem on tõenäosus detekteerida mittesoovitud produkte, näiteks praimerite dimeere (Watson, 1989; Ferrie *et al.*, 1992).

Kuna PCR protsess koosneb paljudest tsüklitest ja amplifikatsiooni efektiivsuse ja lõppsaagise hulga vahel on eksponentsiaalne seos, siis juba väikesed muutused amplifikatsiooni efektiivsuses (näiteks reaktsioonitingimuste optimeerimisest, praimerite järjestusest või muudest faktoritest tingitult) viivad oluliste muutusteni amplifikatsiooni lõppsaagises (produktihulgas), isegi olukorras, kus esialgne sihtmärk DNA molekulide arv on sama (Arezi *et al.*, 2003).

1.2. Amplifikatsiooni efektiivsust mõjutavad tegurid

Lisaks PCR katsetingimustele, mida me käesolevas töös ei käsitle, mõjutavad amplifikatsiooni efektiivsust ka erinevad praimeritega seotud faktorid ja interaktsioonid, mida saab kasutada amplifikatsiooni efektiivsuse ennustamiseks praimerite disaini ja valimise käigus enne PCR katseid.

1.2.1. Praimeri ja DNA järjestuse vahelise hübridisatsiooni tugevus

PCR praimerite ja DNA järjestuse hübridiseerumist kui üheaahelaliste DNA ahelate seondumist kaheaahelaliseks DNA-ks, moodustuva dupleksi stabiilsust (hübridisatsiooni tugevust) ja seda mõjutavaid faktoreid on uuritud küllaltki palju. Stabiilne seondumine (stabiilse dupleksi moodustumine) PCR praimerite ja DNA järjestuse vahel on aluseks polümeraasi seondumisele ja praimerite pikendamisele.

Toatemperatuuril on DNA kaheaahelaline heeliks, milles ahelad püsivad koos tänu Watson-Crick paardumise vesiniksidemetele lämmastikaluste vahel (G ja C aluste vahel kolm ja A ja T vahel kaks sidet) ning aluste vahelisele stäkingule (mittekovalentne interaktsioon sama ahela kõrvuti asetsevate või ka vastasahela lämmastikaluste vahel) (Kool, 1997). Kuna vesiniksidemete energeetiline väärtus on suhteliselt madal, pole vesiniksidemeid stabiilse kaheaahelalise DNA sünteesiks polümeraasi poolt otseselt tarvis, kuid aluste paardumise selektiivsuse tagamiseks on Watson-Crick vesiniksidemete moodustumine oluline (Kool, 2001). Stäkkingu energeetiline väärtus sõltub lahuse koostisest, keskkonna temperatuurist, DNA primaarstruktuurist (stäkkingus olevate nukleotiidide ja neid ümbritsevate nukleotiidide tüübist antud ahelas) (Kool, 1997). Aluste stäking on tugevaim puriinide vahel (nt A-A, G-G) ja nõrgim pürimidiin-pürimidiin vahel (nt T-T); puriin-pürimidiin stäking (nt A-T) on vahepealse tasemega (puriin-pürimidiin \geq pürimidiin-puriin) (Kool, 2001).

DNA ahelate vaheliste dupleksite stabiilsust kirjeldatakse Gibbsi vabaenergia muudu (ΔG) abil. Positiivne ΔG väärtus näitab, et süsteem liigub reagentide tootmise suunas, oligote puhul üheaahelalise seisundi suunas. Negatiivne väärtus näitab, et süsteem liigub produktide tootmise suunas, oligote puhul kaheaahelalise seisundi suunas. ΔG väärtus 0 tähendab kahe oleku tasakaalu seisundit. Mida väiksem on ΔG väärtus, seda stabiilsem on dupleks (SantaLucia, 2007; SantaLucia ja Hicks, 2004).

Üks laialt kasutatav mudel, mis kirjeldab DNA oligonukleotiidide seondumise stabiilsust ja polümeeri termodünaamikat on nn lähima naabri mudel (*Nearest Neighbor*

Model). Nukleiinhapete lähima naabri mudel eeldab, et DNA dupleksi stabiilsus sõltub naabruses olevate aluspaaride identsusest ja suunast. Lähima naabri mudeli kohaselt on DNA dupleksi stabiilsus ja sulamine teatud praimerite ja soolade kontsentratsiooni juures ennustatav kõikide dupleksi koosseisus olevate kõrvuti asetsevate nukleotiidide dimeeride interaktsioonide suhtelise stabiilsuse (Gibbsi vabaenergia) ja temperatuurist sõltuvast käitumisest (entalpia, entroopia) kaudu (Marky ja Breslauer, 1982). Kogu dupleksi Gibbsi vaba energia on üldistatult dupleksi kõikide dimeeride vabaenergiate summa (Breslauer *et al.*, 1986). Aluspaari vabaenergia sisaldab nii aluste vahelise paardumise kui ka stäkingu interaktsioonide energiat (Pienaar *et al.*, 2006).

Oligonukleotiidide dupleksite vabaenergia muudu arvutamisel arvestatakse ka dupleksi moodustumise alustamisega seotud vabaenergia muuduga, võttes arvesse erinevusi dupleksite vahel, millel on terminaaalsed A-T vs terminaaalsed G-C paarid (Allawi ja SantaLucia, 1997; Gray, 1997). Lähima naabri konseptsiooni kohaselt sõltub DNA molekuli vabaenergia lähimatest naabritest ja lahuse soola kontsentratsioonist (SantaLucia, 1998). SantaLucia (1998) leidis, et DNA molekuli pikkuse mõju on seotud soola kontsentratsiooniga, aga mitte DNA järjestusega.

Kõige stabiilsem PCR praimer seondumiskoht on praimeride järjestusega täies pikkuses komplementaarne järjestus, kuid hübriidsatsioon ja stabiilse dupleksi moodustumine praimerite ja sihtmärk-DNA vahel võib toimuda ka teatud hulga mittepaardunud nukleotiidide esinemisel (mittekomplementaarne hübriidsatsioon), moodustades mittepaardunud nukleotiididega dupleksi (Zhao ja Guan, 2010). Erinevate termodünaamika parameetrite väärtused on eksperimentaalselt leitud iga paardunud või mittepaardunud aluspaari jaoks (Allawi ja SantaLucia, 1997; Allawi ja SantaLucia, 1998a; Allawi ja SantaLucia, 1998b; Allawi ja SantaLucia, 1998c; Peyret *et al.*, 1999).

Kuna termodünaamikal põhinevad mudelid on keskendunud vaid DNA ahelate hübriidsatsioonil tekkiva dupleksi stabiilsuse kirjeldamisele ning ei suuda arvestada polümeraasi seondumiseks ja praimerite pikendamiseks vajalike tingimustega, siis ei saa termodünaamika seadustel põhinevaid mudeleid rakendada praimerite pikendamise ja amplifikatsiooni efektiivuse ennustamiseks.

1.2.2. DNA polümeraasi seondumise ja praimerite pikendamise efektiivsus

Lisaks praimerite sihtmärk-DNA järjestusega seondumise edukusele ning moodustuva dupleksi stabiilsusele sõltub amplifikatsiooni efektiivsus ja saagis olulisel määral ka sellest, kui edukalt suudab polümeraas praimerit ja sihtmärk-DNA vahelise dupleksiga seonduda ning praimerit pikendada. Seega ei saa amplifikatsiooni efektiivsuse ennustamisel lähtuda ainult praimerite ja sihtmärkjärjestuste hübridisatsiooni efektiivsusest.

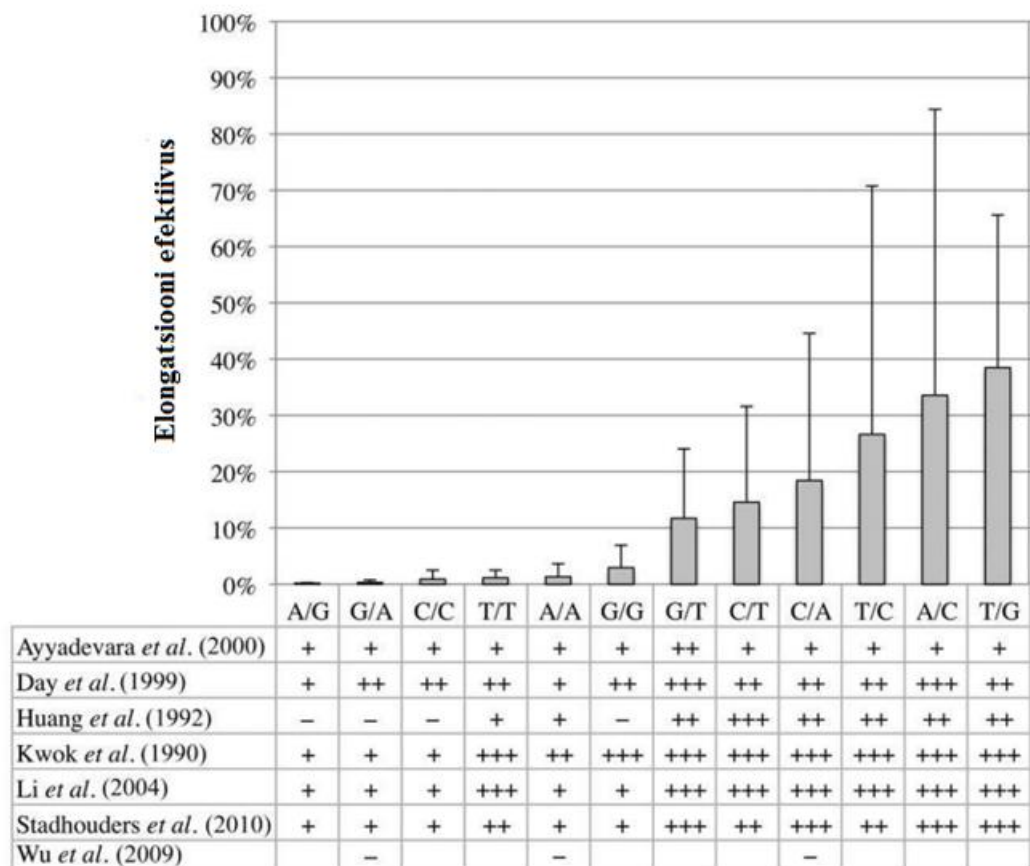
Kui DNA ahelate vahelise (sh praimerit ja sihtmärk-DNA vahelise) hübridisatsiooni efektiivsust on piisavalt põhjalikult kirjeldatud tasakaalulise termodünaamika mudelitega ja üritatud teatud määral rakendada ka praimeridisaini programmides, siis elongatsiooni efektiivsust (primerit pikendamise efektiivsust) on vähem uuritud. Kuna praimerite pikendamine polümeraasi abil on kompleksne protsess, mis lisaks polümeraasi seondumise efektiivsusele dupleksiga sõltub ka nukleotiidide ahelasse lülitamise efektiivsusest, siis ei ole seda seni õnnestunud väga edukalt modelleerida ning rakendada praimerite disainil.

DNA sünteesi käivitamise eelduseks on stabiilne praimerit järjestuse ja sihtmärk-DNA järjestuse vahel moodustunud dupleks. Polümeraas seondub dupleksile ja katalüüsib nukleotiidide lisamist praimerit 3'-OH rühma külge vastavalt komplementaarsusele (Watson-Crick aluspaardumisele, Watson, 1971) sihtmärk-DNA järjestusega. Mittepaardunud nukleotiidid dupleksis mõjutavad lisaks dupleksi enda stabiilsusele ka otseselt polümeraasi seondumist dupleksile. Kuna polümeraas seondub praimerit 3'-otsale, siis on avaldatud mitmeid töid just mittepaardunud 3'-otsa nukleotiidide mõju kohta amplifikatsioonile (Kwok *et al.*, 1990; Huang *et al.*, 1992; Day *et al.*, 1999; Ayyadevara *et al.*, 2000; Wu *et al.*, 2009; Stadhouders *et al.*, 2010). Kui praimerit järjestuse keskel või 5'-otsas paiknev üksik mittepaardunud nukleotiid polümeraasi seondumist üldiselt ei mõjuta, siis praimerit 3'-otsas paiknevad mittepaardunud nukleotiidid mõjutavad oluliselt ka polümeraasi seondumist ja praimerite pikendamist ning seeläbi kogu amplifikatsiooni efektiivsust (Ayyadevara *et al.*, 2000) ning saagist (Piao *et al.*, 2009). Mida praimerit 5'-otsa pool mittepaardunud nukleotiidid paiknevad, seda väiksem on nende mõju praimerite pikendamisele (Wu *et al.*, 2009).

Kuigi üldiselt kehtib reegel, et 3'-otsas paiknevate üksikute nukleotiidide mittepaardumist on võimalik kasutada järjestuste eristamiseks (Cha *et al.*, 1992; Li *et al.*, 2004; Fu *et al.*, 2008), on vähe pööratud tähelepanu nende tõhususe täpsematele kriteeriumitele, probleemidele järjestuste eristamisel teatud olukordades ning mõistlikumale rakendamisele praimeridisainil (Wright *et al.*, 2014). Dupleksi mittepaardunud 3'-otsa nukleotiidide mõju uurimisel peaks arvestama nii mõju praimerite hübridisatsiooni efektiivsusele kui ka

pikendamise efektiivusele või ainult pikendamise efektiivsuse uurimisel hübriidsatsiooni efektiivsuse muutujana kõrvaldama. Kuigi on näidatud, et praimeri 3'-otsas paiknev mittepaardunud nukleotiid võib hoida ära praimeri pikendamist DNA-st sõltuva DNA polümeraasi poolt (Thweatt *et al.*, 1990; Liang ja Pardee, 1992; Sarkar *et al.*, 1990; Ayyadevara *et al.*, 2000), ei tarvitse üks 3'-otsa mittepaardunud nukleotiid olla siiski piisav järjestuste eristamiseks (Stadhouders *et al.*, 2010; Day *et al.*, 1999; Ayyadevara *et al.*, 2000; Kwok *et al.*, 1990).

On leitud, et erinevat tüüpi mittepaardunud (mittekomplementaarsetel) aluspaaridel praimeri ja sihtmärkjärjestuse vahelises dupleksis on erinev mõju amplifikatsiooni efektiivsusele. Puriin/puriin ja pürimidiin/pürimidiin tüüpi mittepaardunud aluspaarid on üldiselt tugevama mõjuga (võrreldes puriin/pürimidiin või pürimidiin/puriin tüüpi mittepaardunud aluspaaridega) ning põhjustavad madalamat amplifikatsiooni efektiivsust ning amplifikatsiooni lõppsaagist (Lefever *et al.*, 2013). On leitud, et praimerite pikendamisele teadaolevalt kõige tugevamat mõju avaldavad mittepaardunud aluspaarid A/G (A nukleotiid praimeri järjestuses on kohakuti G nukleotiidiga sihtmärkjärjestuses), G/A, C/C ja A/A ning nõrka mõju avaldavad mittepaardunud aluspaarid G/T, C/T, A/C ja T/G. Sümmeetriliste mittepaardunud aluspaaride (nt T/G versus G/T) mõju on omavahel väga sarnane (Stadhouders *et al.*, 2010; Wright *et al.*, 2014; Huang *et al.*, 1992; Ayyadevara *et al.*, 2000). Samas on erinevad töögrupid saanud konkreetset tüüpi 3'-otsa mittepaardunud aluspaaride mõju kohta praimerite pikendamise efektiivsusele väga erinevaid ja ka vastukäivaid tulemusi, mistõttu on raske teha üheseid järeldusi erinevat tüüpi 3'-otsa mittepaardunud aluspaaride erineva mõju kohta. Näiteks on erinevad uuringud (Huang *et al.*, 1992; Christopherson *et al.*, 1997; Day *et al.*, 1999; Ayyadevara *et al.*, 2000) välja toonud väga erinevaid hinnanguid praimeri ja sihtmärkjärjestuse üksiku praimeri 3'-otsa mittepaardunud aluspaari A/C mõju kohta, alates sellest, et mõju praktiliselt puudub ning amplifikatsioon toimub sama edukalt kui Watson-Crick paardumisega A/T puhul (Day *et al.*, 1999) kuni peaaegu täieliku amplifikatsiooni inhibitsioonini (Ayyadevara *et al.*, 2000). Sarnaselt eelnevaga on saadud vastukäivaid tulemusi terminaalse C/A mõju kohta amplifikatsioonile, alates vähesest või täielikult puudevast amplifikatsioonist kuni vaid mõne tsüklilise viiteni (Ayyadevara *et al.*, 2000; Wu *et al.*, 2009; Stadhouders *et al.*, 2010). Üks kõige tõenäolisem seniste uuringute varieeruvate tulemuste (joonis 1) põhjus on tõenäoliselt erinevate töögruppide poolt kasutatud erinevad katsetingimused (nt seondumistemperatuur, praimeri pikkused jne).



Joonis 1. Praimerite pikendamise (elongatsiooni) efektiivsuste võrdlus erinevate avaldatud uuringute ja erinevat tüüpi 3'-otsa mittepaardunud aluspaaride korral. Graafikul on konkreetses uuringus (Wright *et al.*, 2014) leitud keskmised elongatsiooni efektiivsused kõikide võimalike mittepaardunud (praimer/sihtmärk-DNA) 3'-otsa aluspaaride puhul ja alumises tabelis varem avaldatud uurimuste tulemused (- amplifikatsioon puudub, + nõrk amplifikatsioon, ++ mõõdukas amplifikatsioon, +++ tugev amplifikatsioon). Vea piirid (*error bars*) näitavad standardhälvet erinevate testitud sama tüüpi 3'-otsa mittepaarduva aluspaariga praimerite keskmiste elongatsiooni efektiivsuste vahel (Wright *et al.*, 2014).

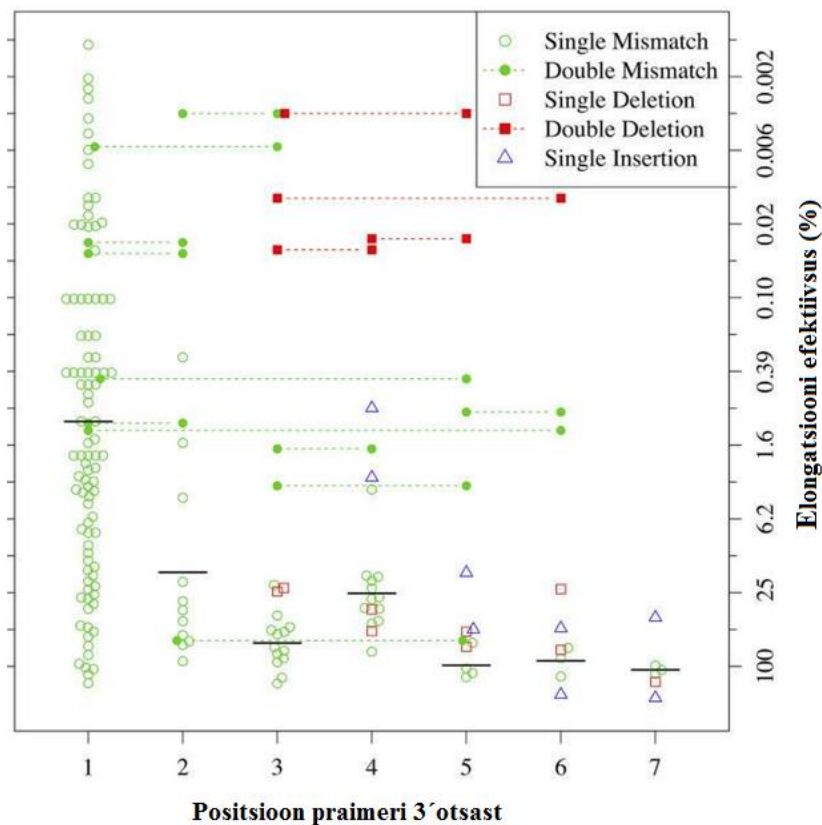
Üksiku praimeri ja sihtmärkjärjestuse mittepaardunud nukleotiidi mõju amplifikatsioonile sõltub selle asukohast. On näidatud, et mida rohkem 3'-otsa lähedal (kuni 4. - 5. positsioon) see paikneb, seda suuremat mõju elongatsiooni efektiivsusele avaldab. Üksikud mittepaardunud nukleotiidid kaugemal kui 5. positsioon elongatsiooni efektiivsust oluliselt ei mõjuta (Zhao ja Guan, 2010; Stadhouders *et al.*, 2010) ning amplifikatsiooni toimumist ei blokeeri (Lefever *et al.*, 2013). Ühe mittepaardunud nukleotiidi poolt põhjustatud inhibitsioon on kõige märgatavam 3'-otsas paikneva (1. positsiooni) mittepaardunud nukleotiidi puhul. On näidatud, et kui üksik mittepaardunud nukleotiid paikneb 3'-otsast kaugemal kui 7. positsioon,

siis on selle mõju peaaegu olematu (Lefever *et al.*, 2013). Üksikute 3'-otsa lähedal paiknevate indelite (insertsioonide/deletsioonide) mõju on sarnane üksikute mittepaardunud nukleotiidide mõjule samades positsioonides (Wright *et al.*, 2014).

Saadud tulemused mittepaardunud nukleotiidide asukoha olulisuse kohta langevad kokku teadmistega DNA polümeraasi seondumiseks ja praimerite pikendamiseks vajaliku nukleotiidi arvu kohta, mis on teadaolevalt 5 - 8 nukleotiidi (Johnson and Beese, 2004; Zhao ja Guan, 2010; Catalano *et al.*, 1990; Li *et al.*, 1998). Seetõttu mittepaardunud nukleotiidid praimer 3'-otsast kuni 5 - 8 bp (aluspaari) kaugusel võivad avaldada suuremat mõju amplifikatsioonile, mõjutades lisaks praimer ja sihtmärkjärjestuse dupleksi stabiilsuse oluliselt ka polümeraasi seondumise ja praimerite pikendamise edukust.

On näidatud, et üksikul mittepaardunud nukleotiidil praimer 3'-otsas võib olla mõju PCR efektiivsusele vaid esimestes PCR tsüklites ning see ei mõjuta või mõjutab PCR lõppsaagist (molekulide arv reaktsiooni lõpus) väga vähe. Kui mittepaardunud nukleotiid on lülitatud juba ahelasse, siis toimub PCR edasi tavapärase efektiivsusega (Huang *et al.*, 1992; Boyle *et al.*, 2009).

Kui ühe 3'-otsa lähedal paikneva mittepaardunud aluspaari mõju amplifikatsioonile on uuritud küllaltki palju, siis suurema arvu mittepaardunud aluspaaride mõju sõltuvalt nende arvust, asukohast, tüübist on uuritud tunduvalt vähem. On leitud, et kahe mittepaardunud aluspaari mõju on oluline ka 4. positsioonist (primer 3'-otsast) kaugemal ning mõju on suurem võrreldes üksiku mittepaardunud aluspaari mõjuga amplifikatsioonile. Kui Stadhouders *et al.*, (2010) uuringu tulemused näitasid, et kahe mittepaardunud aluspaariga dupleksi puhul primeri pikendamist ei toimu, siis (Wright *et al.*, 2014) tulemused (joonis 2) tuvastasid teatud tüüpi kahe mittepaardunud aluspaari kombinatsioonide puhul siiski amplifikatsiooni. Erinevad tulemused on seletatavad sellega, et Stadhouders *et al.* (2010) puhul kasutati katsetes sulamistemperatuurile lähedasi seondumistemperatuure, mistõttu peegeldavad tulemused mittepaardunud aluspaaride mõju nii primerite pikendamise efektiivsusele kui ka hübridisatsiooni efektiivsusele, kuid Wright *et al.* (2014) on oma uurimuses hübridisatsiooni efektiivsuse maksimaliseerinud madala seondumistemperatuuriga ning tulemused peegeldavad mittepaardunud aluspaaride mõju ainult primerite pikendamise efektiivsusele. Kahe mittepaardunud nukleotiidi esinemisel primeri järjestuses, millest üks paikneb primer 3'-otsas ja teine 3'-otsast kuni 5 bp kaugusel, on amplifikatsiooni protsess oluliselt pidurdatud (Lefever *et al.*, 2013).



Joonis 2. Elongatsiooni efektiivsuse sõltuvus ühe ja kahe mittepaarduva aluspaari, deletsiooni ja inserstiooni positsioonist. Positsioon 1 on praimer 3'-ots, positsioon 2 on 3'-otsast järgmine positsioon ja nii edasi. Deletsioonide ja inserstioonide puhul tähistab positsiooni number positsiooni, kust on nukleotiid on kadunud või juurde tulnud võrreldes täiskomplementaarse praimer järjestusega. Horisontaalne pidevjoon tähistab iga positsiooni üksikute mittepaarduvate aluspaaride keskmist elongatsiooni efektiivsuse väärtust. Kahe mittepaarduva aluspaariga või deletsiooniga positsiooni samas praimeris on ühendatud katkendjoonega. Hübridisatsiooni efektiivsus on maksimeeritud, kasutades madalat seondumistemperatuuri (Wright *et al.*, 2014).

Selleks, et kontrollida erinevas positsioonis (sh 3'-otsas) paikneva ühe või kahe mittepaardunud aluspaari mõju meie poolt kasutatavatel katsetingimustel ja arvestada seda käesolevas töös välja töötatud praimerite seondumise ja pikendamise efektiivsust hindava skoori väljatöötamisel, tehti antud töö raames ka ise mõned PCR katsed.

On leitud, et mida rohkem mittepaardunud aluspaare ja mida rohkem 3'-otsa lähedal need paiknevad, seda suurem on nende mõju amplifikatsiooni efektiivsusele (Lefever *et al.*, 2013). Nelja mittepaardunud nukleotiidi olemasolu ühes praimeris võib amplifikatsiooni peatada peaaegu täielikult (va juhul, kui on neli järjestikust mittepaardunud nukleotiidi praimer 5'-otsas). Samuti on leitud, et kuus või rohkem mittepaardunud nukleotiidi pärisuunalises ja

vastassuunalises praimeris kokku, juhul kui ühel praimeril on vähemalt kolm mittepaardunud nukleotiidi, blokeerib amplifikatsiooni täielikult (Lefever *et al.*, 2013).

Praimeri 3'-otsas paikneva mittepaardunud nukleotiidi pikendamise kiirus sõltub DNA polümeraasi võimest erinevat tüüpi mittepaardunud aluspaare sisaldava dupleksiga seonduda ja praimerit pikendada. Praimeri pikendamisel seondub ensüüm praimer-DNA dupleksiga praimer 3'-otsas, mille järel hakkab DNA-ensüümi kompleks siduma dNTP-sid ja lülitama neid DNA ahelasse (Fisher ja Korn, 1981; Detera *et al.*, 1981). Seega võib mittepaardunud 3'-otsa nukleotiididega praimer 3'-otsa pikendamise kiiruse vähenemine olla põhjustatud kas polümeraasi ja DNA kompleksi seondumiskonstandi väärtuse langusest või seondunud ensüümi väiksemast mittepaardunud nukleotiididega praimerite pikendamise kiirusest või nende kahe faktori kombinatsioonist (Huang *et al.*, 1992).

Taq DNA polümeraasil puudub 3'→5' eksonukleasne aktiivsus, mis parandaks praimer 3'-otsas paiknevaid mittepaardumisi ning selle omaduse puudmine mõjutab oluliselt 3'-otsas paiknevate mittepaardunud nukleotiidide pikendamise efektiivsust ja seeläbi kogu amplifikatsiooni efektiivsust (Ayyadevara *et al.*, 2000). On näidatud, et Taq polümeraas seondub sarnase afiinsusega 3'-otsa mittepaardunud aluspaare sisaldavate või ainult paardunud aluspaare sisaldava praimer-DNA dupleksiga, kuid pikendab 3'-otsa mittepaardunud aluspaare sisaldavat dupleksit tunduvalt vähem efektiivselt võrreldes teiste ensüümidega. Seetõttu on Taq polümeraas (mille puudub 3'→5' eksonukleasne aktiivsus) sobiv rakendusteks, kus on vaja PCR abil eristada üksikute nukleotiidide poolest erinevaid DNA järjestusi (nn alleelspetsiifiline amplifikatsioon) ning on oluline, et polümeraas ei oleks võimeline eemaldama mittepaardunud nukleotiide ja kärbitud praimerit pikendama (Huang *et al.*, 1992; Roux, 2009).

1.3. Taksoni-spetsiifiliste testide väljatöötamine

Üha enam kasutatakse molekulaarseid meetodeid organismide tuvastamiseks. PCR on usaldusväärne, odav, lihtsa tööpõhimõttega ja tundlik meetod, mis võimaldab spetsiifiliste praimerite abil amplifitseerida ühe ja konkreetse taksoni DNA-d ning selle kaudu tuvastada vastava taksoni (nt liigi, perekonna, sugukonna) olemasolu. Seetõttu kasutatakse PCR-il põhinevaid teste laialdaselt väga erinevates valdkondades: mikrobioloogias, toiduainete koostise analüüsil, biomeditsiinilistes uuringutes, põllumajanduses ja keskkonnaproovide analüüsis (liigilise mitmekesisuse uurimiseks) (Oggioni *et al.*, 2002; Elfnifro *et al.*, 2000; Haas *et al.*, 1998; Duitama *et al.*, 2009; Croning *et al.*, 2010; Yang *et al.*, 2012; Bryksin and Matsumura, 2010; Boom *et al.*, 2002; Van Borm and Boomsma, 2002; Jarman *et al.*, 2002; Jarman *et al.*, 2003; López-Calleja *et al.*, 2005; Wilcox *et al.*, 2013). Edaspidi on vastavaid teste nimetatud lihtsuse mõttes taksoni-spetsiifilisteks testideks, mis on disainitud, et tuvastada erinevate fülogeneetiliste tasemete taksonideid (nt sugukond, perekond, liik, alamliik, tüvi).

Taksoni-spetsiifiliste diagnostiliste testide puhul kasutatakse amplifikatsiooniks tihti DNA-proovi, mis sisaldab paljude erinevate (sh fülogeneetiliselt lähedaste ja seetõttu sarnase DNA-järjestustega) organismide DNA-d (nt toiduainest või keskkonnaproovist eraldatud DNA). Näiteks allergeensete (allergeenseid ühendeid sisaldavate) taimede tuvastamisel toidust võib olla vajalik eristada sama sugukonna allergeenseid taimeliike pekani-hikkoripuu (*Carya illinoensis*) ja kreeka pähklikpuu (*Juglans regia*), sama perekonna (nt *Brassica*) liike või koguni sama liigi alamliike (perekonna *Triticum* liigid ja alamliigid), mille DNA-d esinevad sageli amplifitseeritavas DNA proovis koos. Sihtmärk-DNA on konkreetse taksoni DNA, mida soovitakse amplifikatsioonil tuvastada. Mittesihtmärk-DNA on analüüsitavas DNA-proovis koos sihtmärk-DNA-ga esinev teiste organismide (sh fülogeneetiliselt lähedaste taksonite) DNA, mida ei soovita testiga tuvastada. Konkreetne mittesihtmärk-taksonite komplekt võib sõltuvalt sihtmärk-taksonist ja testi kasutamise eesmärgist (analüüsitavast proovimaterjalist) olla väga erinev.

PCR-il tugineva taksoni-spetsiifilise testi aluseks on taksoni-spetsiifilised praimerid, mis amplifitseerivad erinevate taksonite DNA-de segust vaid sihtmärk-taksoni DNA-d. Sõltuvalt organismist, mida tuvastada soovitakse, võib taksoni-spetsiifiliste praimerite disain olla aga väga keerukas ja ajamahukas. Spetsiifiliste praimerite disaini protsess hõlmab endas mitte ainult oligonukleotiidide genereerimist sihtmärkjärjestuse alusel, vaid ka disainiks sobiva geenipiirkonna valikut, DNA järjestuse poolest sarnaste ja PCR-il soovimatute mittesihtmärk-

taksonite määratlemist, sihtmärk- ja mittesihtmärk-taksonite DNA järjestuste hankimist andmebaasidest või sekveneerimise abil ja praimerite spetsiifilisuse kontrolli *in silico*.

Sobiv geenipiirkond on DNA järjestus, milles leidub piisavalt palju tuvastatava taksoni DNA järjestuste vahel mittevarieeruvaid (et tagada piisav tundlikkus sihtmärkjärjestuste tuvastamisel ja vältida valenegatiivset tulemust) ja samal ajal tuvastatavate ja mittetuvastatavate taksonite vahel varieeruvaid (et vältida samaaegset mittesihtmärk-taksoni amplifikatsiooni ja vältida valepositiivseid tulemusi) kohti (Villard ja Malausa, 2013). Et sihtmärk-taksoni ja sarnase DNA järjestusega mittesihtmärk-taksonite DNA järjestusi võrrelda ja leida sobivad kohad spetsiifiliste praimerite disainiks, on oluline, et kasutatava regiooni järjestusi oleks andmebaasides piisavalt palju nii sihtmärk-taksoni kui ka erinevate lähedaste mittesihtmärk-taksonite jaoks.

Vaatamata sellele, et sekveneerimine on muutumas järjest odavamaks ja kättesaadavamaks ning sekveneeritud on päris paljude organismide täisgenoomid, on veel terve hulk organismirühmi (näiteks taimed), mille täisgenoome on sekveneeritud veel liiga vähe, et neid taksoni-spetsiifiliste praimerite disainiks kasutada. Sellisel juhul on võimalik disainiks kasutada ainult andmebaasides olevate piiratud hulga üksikute sekveneeritud geenide genoomseid järjestusi. Väga paljude geenide puhul on andmebaasides ainult geenide mRNA järjestused. Kasutades disainiks mRNA järjestusi, mille ekson-intron piirid ei ole teada ning ei ole ennustatavad teiste genoomsete järjestuste abil, on oht, et disainitavad PCR praimerid satuvad ekson-intron piiridele ning hiljem PCR-i käigus genoomset DNA-d ei amplifitseerita. Lisaks tuleb sobiva geeniregiooni valikul arvestada erinevate genoomide (kloroplasti, mitokondri ja tuuma genoom) ja geeniregioonide eripäradega, geenide koopiaarvuga, järjestuse varieeruvusega jne. Kuigi praimerite disainiks sobiva geenipiirkonna valimine võib olla ajakulukas töö, on seda etappi keerukas täielikult automatiseerida.

Taksoni-spetsiifilise testi jaoks praimerite disain hõlmab endas ka mittesihtmärk-taksonite määratlemist ja järjestuste hankimist andmebaasidest. Mittesihtmärk-järjestusteks konkreetse testi puhul on eelkõige sihtmärk-taksoni DNA järjestustega homoloogsed DNA järjestused, mis võivad esineda analüüsitavad proovis koos sihtmärk-järjestustega. Mittesihtmärk-järjestusi on andmebaasidest võimalik leida homoloogiaotsingu programmide abil (nt BLAST, Altschul *et al.*, 1997).

Sõltuvalt sihtmärk-taksonist ja konkreetsest geenipiirkonnast võib tuvastatava sihtmärk-taksoni konkreetse geenipiirkonna järjestusi andmebaasis olla väga palju (kümneid, sadu, tuhandeid või koguni kümneid tuhandeid), mis võivad rohkemal või vähemal määral

varieeruda. Kuna mittepaarduvad aluspaarid praimerite ja DNA järjestuse vahel võivad vähendada nii duplexi stabiilsust kui praimerite pikendamise efektiivsust, tuleks võimalusel vältida praimerite disaini geenipiirkonda, mille järjestused sihtmärk-taksoni siseselt varieeruvad (või vajadusel lubades varieeruvaid positsioone ainult praimerite 5'-otsa lähedal). Kuna samaaegselt tuleb arvestada nii sihtmärk-taksoni järjestuste vahelist varieerumist kui ka sihtmärk-taksoni ja mittesihhtmärk-taksoni järjestuste vahelist varieerumist (mis võib fülogeneetiliselt lähedaste liikide puhul olla väga väike), siis võib taksoni-spetsiifilise testi jaoks disainitavate praimerite jaoks sobivate kohtade leidmine olla suur väljakutse.

Kuigi PCR-põhised diagnostikatestid taksonite tuvastamiseks ja identifitseerimiseks on väga laialt kasutusel erinevates valdkondades, ei ole väga palju arvutiprogramme, mille abil disainida spetsiifilisi primereid või hinnata disainitud praimeripaare sobivust konkreetse taksoni-spetsiifilise DNA testi jaoks. Arvutiprogrammide abil automatiseeritud taksoni-spetsiifiliste praimerite disaini etapid lihtsustaksid ja suurendaksid veelgi PCR-il baseeruvate tuvastamismeetodite kasutamist ja efektiivsust.

1.4. Olemasolev tarkvara praimerite spetsiifilisuse kontrollimiseks

PCR primereid, mis amplifitseeriksid teatud DNA regiooni, võib üldjuhul paigutada väga paljudesse kohtadesse sihtmärk-järjestusel. On terve rida programme, mis aitavad leida praimerite seondumiseks kõige sobivamaid ja soovitud produkti amplifitseerivaid kohti järjestusel (nt Untergasser *et al.*, 2012; Nielsen ja Knutsen, 2002). Enamik neist programmidest on keskendunud praimerite disainile, mis võimaldavad amplifitseerida DNA regiooni ühest kindlast järjestusest (Primer3, Rozen ja Skaletsky, 2000; PerlPrimer, Marshall, 2004 jne), aidates kasutajal disainida ja valida erinevatele biokeemilistele jm kriteeriumitele kõige paremini vastavaid primereid. Samas et aita need programmid kasutajal disainida või valida kõige paremaid praimeripaare taksoni-spetsiifilise testi jaoks.

Primer3 (Rozen ja Skaletsky, 2000; Untergasser *et al.*, 2007; Untergasser *et al.*, 2012) on arvutitarkvara, mis on praimeridisainiks laialt kasutusel. Programm disainib konkreetset sihtmärkjärjestust amplifitseerivad praimeripaarid, mis on kontrollitud võimalike konkureerivate mittespetsiifiliste seondumiste suhtes iseenda (sekundaarstruktuurid) ja teiste praimeritega (dimeerid). Primer3 väljundfailis on praimerid esitatud pingereas, mis tugineb parameetriväärtuste erinevusele määratud optimumväärtusest. Primer3 ei kontrolli primerite mittespetsiifilisi seondumisi mittesihhtmärk-taksonite järjestustele ning kui kasutaja on liigselt

piiranud väljundina saadavate praimerite hulka, võib võimalik spetsiifiline praimeripaar jääda ka väljundfailist välja.

Taksoni-spetsiifilise testi jaoks disainitud praimeripaar peaks ideaalsel juhul amplifitseerima ainult defineeritud sihtmärk-taksoni DNA järjestusi ja ei tohiks amplifitseerida mitte ühtegi samas analüüsitavas proovis esinevat mittesihtmärk-taksoni DNA järjestust. Selleks, et kasutaja saaks kontrollida või hinnata disainitud praimerite unikaalsust (spetsiifilisust), hinnata praimeripaari mittespetsiifiliste seondumiste ja amplifikatsiooni efektiivsust, tuleb tal tihti kasutada selleks eraldi programme.

On rida programme praimerite mittespetsiifilisi seondumiskohtade leidmiseks konkreetse genoomijärjestuse piires (GenomeTester, Andreson *et al.*, 2006; PRIMEX, Lexa ja Valle, 2003; BiSearch, Aranyi *et al.*, 2006), kuid oluliselt vähem programme, mis võimaldavad leida mittespetsiifilisi seondumiskohti ja hinnata praimerite pikendamise efektiivsust erinevate mittesihtmärk-taksonite järjestuste puhul.

1.4.1. Tarkvara praimerite mittespetsiifiliste seondumiste ennustamiseks

Praimeridisaini programmid PerlPrimer (Marshall, 2004) ja Primer3Plus (Untergasser *et al.*, 2007) soovivad kasutajal praimerite mittespetsiifiliste seondumiste leidmiseks sihtmärk- või mittesihtmärkjärjestustelt kasutada USA Riikliku Biotehnoloogilise Informatsiooni Keskuse (NCBI, *National Center for Biotechnology Information*) BLAST (*Basic Local Alignment Search Tool*) otsingut.

BLAST on homoloogiaotsinguteks laialt kasutatav programm (Xu *et al.*, 2002; Weckx *et al.*, 2005; Rouchka *et al.*, 2005), mis otsib uuritava järjestusega (milleks võib olla ka näiteks praimer) sarnaseid järjestusi andmebaasist (Altschul *et al.*, 1997). BLAST algoritmi kasutavad mitmed programmid praimerite seondumiskohtade ennustamiseks (nt Primer-BLAST, Ye *et al.*, 2012; QuantPrime, Arvidsson *et al.*, 2008; PRIMEGENS, Xu *et al.*, 2002). Kõik tööriistad, mis kasutavad otsingu-algoritmina BLAST-i, omavad minimaalsest sõna pikkusest tulenevat piirangut. Näiteks BLAST otsing, mis kasutab minimaalset sõnapikkust 7 (nt Primer-BLAST puhul) ei leia järjestusi, millel on võrreldes otsingusõnaga (näiteks praimerijärjestusega) 6 või vähem järjestikust komplementaarset paardumist. Seega näiteks kasutades otsingusõnana 20 nukleotiidi pikkust praimerijärjestust, milles on 7. ja 14. positsioonis konkreetse (mitte)sihtmärk-DNA järjestusega mittepaarduvad nukleotiidid, jääb selline

(mitte)sihtmärk-järjestus BLAST otsinguga tuvastamata, kuigi praimeris on vaid 2 mittepaardunud nukleotiidi, mis üldjuhul praimerite seondumist ja pikendamist ei takista.

BLAST programm ei ole praimerite spetsiifilisuse kontrollimiseks väga hea tööriist ka seetõttu, et see kasutab lokaalse joonduse algoritmi ning seetõttu ei võimalda kokku lugeda täpset mittepaardunud aluspaaride arvu kahe võrreldava järjestuse (primeri ja sihtmärk-järjestus) vahel (Ye *et al.*, 2012; Altschul *et al.*, 1997). Lisaks eelnevale ei aita BLAST omavahel võrrelda ja sorteerida primereid või praimeripaare spetsiifilisuse alusel. Olukorras, kus analüüsitavaid primereid või praimeripaare on väga palju, kasutatavad andmehulgad on suured, võib sobivaima praimeripaari valimine olla väga ajakulukas protsess.

MFEprimer-2.0 on üks väheseid termodünaamika seadustel (ja lähima naabri mudelil) põhinevaid programme primeri ja DNA-järjestuse vahelise seondumise stabiilsuse hindamiseks (Qu *et al.*, 2012). Programm tugineb teadmisel, et eduka PCR reaktsiooni aluseks on primeri 3'-otsa järjestuse unikaalsus (mis tagab spetsiifilisuse) ning kogu primeri järjestus (dupleksi stabiilsuse tagamiseks). Kasutaja annab ette praimeripaari ja programm otsib andmebaasi (peamiselt ainult mudelorganismide DNA/RNA järjestusi sisaldav) järjestustelt primeri 3'-otsaga (kuni 5 - 9 nukleotiidi) stabiilseid duplekseid moodustavaid primeri seondumissaite, hindab dupleksite stabiilsust kogu primeri pikkuse ulatuses ning lõpuks sorteerib amplikoni suuruse jm parameetriväärtuste alusel välja ainult eeldatavalt amplifikatsiooni andvad seondumised ja ennustatud amplikonid. Programm arvutab stabiilsuse hindamiseks iga seondumise Gibbsi vabaenergia (ΔG) ja sulamistemperatuuri (T_m) väärtused. Antud programmi puhul ei saa kasutaja ise sihtmärkjärjestusi ja mittesihhtmärk-järjestusi ette anda ning etteantud andmebaasides on kokku ainult umbes paarkümmend erinevatest fülogeneesiharudest pärineva liigi genoomset järjestust, millest ei ole abi praimerite seondumise hindamiseks, praimerite disainil lähedaste taksonite eristamiseks. Samuti ei aita programm kasutajal sorteerida ja valida taksoni-spetsiifilise testi jaoks kõige sobivamat praimeripaari.

1.4.2. Tarkvara mittespetsiifilise amplifikatsiooni ennustamiseks

Lisaks programmidele, mis ennustavad pärisuunalise ja vastassuunalise primeri seondumisi, on loodud ka programme, mis üritavad ennustada praimeripaari mittespetsiifilise amplifikatsiooni toimumist või mittetoimumist.

Enamik loodud programmidest põhineb BLAST algoritmil (nt Primique, Fredslund ja Lange, 2007; QuantPrime, Arvidsson, 2008; PRIMEGENS, Xu *et al.*, 2002, Srivastava *et al.*, 2008; Primer-Blast, Ye *et al.*, 2012; PRISE2, Huang *et al.*, 2014) ning seetõttu omavad mitmeid BLAST algoritmist tulenevaid piiranguid võimalike seundumiskohtade leidmisel. Mitmed neist programmidest võimaldavad leitud seundumissaidi puhul teatud määral arvestada ka praimeride 3' otsa mittepaarduvate nukleotiidide mõju amplifikatsioonile (nt PRISE2, Huang *et al.*, 2014; Primique, Fredslund ja Lange, 2007; Primer-BLAST, Ye *et al.*, 2012), kuid enamasti eeldades, et kolm mittepaardunud aluspaari praimeride 3'-otsa lähedal on piisav, et amplifikatsiooni ära hoida ning kaugemal paiknevate mittepaarduvate aluspaaride puhul positsioonist sõltuvat mõju ei arvestada.

Mitmed programmid võrdlevad järjestusi ning otsivad sihtmärk- ja mittesihhtmärkjärjestuste vahel sarnase või identse järjestusega kohti kui potentsiaalseid mittespetsiifilise praimerite seundumiskohti, kasutades järjestuste joondust (nt SP-Designer, Villard ja Malausa, 2013). Vead joonduses (tingituna näiteks insertioonidest või deletsioonidest tingitud erinevustest järjestustes, liiga varieeruvatest kohtadest järjestustes ning joondamisprogrammide võimetusest vastavaid regioone joondada) võivad mõjutada sarnase järjestusega regioonide leidmist.

On loodud ka mõned üksikud termodünaamika seadustel tuginevad programmid praimeripaari mittespetsiifilise amplifikatsiooni ennustamiseks, kuid need ei ole kas vabavarana kättesaadavad (nt Thermoblast, <http://www.dnasoftware.com/problems-we-solve/thermoblast>) või sisaldab programmi poolt kasutatav mittesihhtmärk-järjestuste andmebaas vaid piiratud hulga organismide järjestustest (MFEprimer-2.0, Qu *et al.*, 2012),.

Konkreetselt taksoni-spetsiifilise testi jaoks sobivate praimerite puhul on oluline kontrollida ja hinnata praimerite seundumise ja võimaliku pikendamise efektiivsust konkreetsete mitte-sihhtmärk-taksonite järjestuste vastu. Seetõttu on oluline, et programm võimaldaks mittesihhtmärk-järjestusi kasutajal endal sisestada. Paljude seni loodud programmide puhul seda teha ei saa (In-Silico PCR, <http://genome.ucsc.edu/cgi-bin/hgPcr?command=start>; MFEprimer-2.0, Qu *et al.* 2012) või on mittesihhtmärk-järjestuste määramise põhjendamatult ajamahukas (SP-Designer, Villard ja Malausa, 2013; PRISE2, Huang *et al.*, 2014).

Kui analüüsitavaid praimereid on palju, on mõistlik seda teha käsurea programmide abil Linux/Unix keskkonnas. Enamikul programmidest on olemas vaid veebipõhised versioonid

(Primer-BLAST, Ye *et al.*, 2012; QuantPrime, Arvidsson, 2008; PRIMEGENS, Xu *et al.*, 2002, Srivastava *et al.*, 2008; Primique, Fredslund ja Lange, 2007).

Seega on praeguseks loodud mitmeid arvutiprogramme, mis aitavad disainida teatud kriteeriumite alusel spetsiifilisi praimerid, ennustada praimerite (mitte)spetsiifilisi seondumisi või amplifikatsiooni (tuginedes järjestuste sarnasusele, mittepaardunud aluspaaride arvule sihtmärkjärjestuse ja mittesihtmärkjärjestuse vahel või termodünaamika seadustele), kuid mille kasutamisel on mitmeid piiranguid või tuleb lõplik töömahukas praimerite spetsiifilisuse alusel sorteerimine teha kasutajal käsitsi.

Sageli ei ole spetsiifiliste praimerite disainil pudelikaelaks mitte niivõrd programmide kiirus (kui just ei tegeleta täisgenoomijärjestustega), vaid praimerite maksimaalse spetsiifilisuse tagamine (eriti väga lähedaste liikide eristamisel).

Enamik praeguseks loodud programmidest ei suuda leida kõiki võimalikke praimerid (mittespetsiifilisi) seondumissaite (nt BLAST-i algoritmi piirangutest tulenevalt), ei arvesta praimerid ja DNA-järjestuse vahelise dupleksi erinevates positsioonides paiknevate mittepaarduvate aluspaaride mõju praimerite pikendamisele või ei anna praimerite seondumise ja pikendamise efektiivsusele arvulist hinnangut, mida kasutada disainitud praimeritest taksoni-spetsiifilise testi jaoks kõige sobivamate praimeripaaride välja valimiseks. Seetõttu otsustati käesolevas töös luua skoor ning sellel baseeruv arvutiprogramm, mis aitaks hinnata praimerite mittespetsiifilise seondumise ja pikendamise efektiivsust tuginedes mittepaarduvate aluspaaride arvule ja asukohale praimerid ja mittesihtmärkjärjestuse vahelises dupleksis, et sorteerida disainitud praimeripaare ja leida nende seast kõige sobivam praimeripaar konkreetse taksoni-spetsiifilise testi jaoks.

2. Eksperimentaalne osa

2.1. Töö eesmärgid

Haridus- ja Teadusministeeriumi grantidest ja Euroopa Liidu Euroopa Regionaalarengu Fondi programmidest rahastatava projekti „Allergofood“ (2012. – 2015. a.) üheks eesmärgiks oli töötada välja biotehnoloogilised testid allergeenide tuvastamiseks töödeldud toiduainetest ja tootmisliinidelt. DNA-põhiste testide väljatöötamisel tekkis vajadus disainida liigi- või perekonnaspetsiifilised PCR praimerid allergeensete taimede tuvastamiseks toidust.

Taksonite tuvastamisel (näiteks lähedaste liikide eristamisel) PCR abil toidu- või hügieeniproovist tuleb tihti silmitsi seista väljakutsega, et PCR praimerid peavad efektiivselt amplifitseerima sihtmärk-DNA järjestusi, kuid samas ei tohi amplifitseerida samas proovis sisalduvaid väga sarnase, ainult üksikute nukleotiidide poolest erinevaid, mittesihhtmärk-DNA järjestusi. Sellisel juhul on oluline teada, kui palju minimaalselt ja mis positsioonis mittepaardunud aluspaare peaks praimer ja mittesihhtmärkjärjestuse vahel olema, et tagada spetsiifiline ainult sihtmärk-DNA amplifikatsiooni toimumine.

Kuna puudus metoodika ja arvutitarkvara taksoni-spetsiifilise testi jaoks sobivate PCR praimerite disainiks ja välja valimiseks, siis oli vajadus need projekti raames välja töötada.

Antud töö peamiseks eesmärgideks on

- * leida praimerite pikendamise efektiivsust iseloomustav skoor, mis tugineb mittepaardunud aluspaaride arvul ja asukohal praimer ja mittesihhtmärkjärjestuse vahelises dupleksis,
- * kontrollida välja töötatud skoori väärtuse ja amplifikatsiooni toimumise vahelist seost meie katsetingimustes,
- * luua metoodika ja arvutitarkvara taksoni-spetsiifilise testi jaoks sobivate PCR praimerite disainiks ja sobivaima praimeripaari valimiseks tuginedes välja töötatud skoorile,
- * testida loodud metoodikat DNA testi väljatöötamisel allergeensete viljadega taimeliigi pekani-hikkoripuu (*Carya illinoensis*) tuvastamiseks toidust.

2.2. Materjal ja metoodika

2.2.1. Mittepaarduvate nukleotiididega praimerid

Tuginedes kirjanduse andmetele mittepaardunud aluspaaride arvu ja asukoha mõju kohta amplifikatsiooni efektiivsusele, loodi käesolevas töös skoor NAS (*No Amplification Score*), mis iseloomustab praimeri seondumise ja pikendamise efektiivsust konkreetse seondumissaidi puhul. Mida väiksem on praimeri skoori väärtus, seda efektiivsem on eeldatavalt praimeri seondumine ja pikendamine konkreetse seondumissaidi puhul.

Et uurida üksikute, erinevas positsioonis olevate, mittepaarduvate aluspaaride mõju amplifikatsiooni saagisele meie katsetingimustes ning kontrollida käesolevas töös leitud praimerite NAS väärtuste ja amplifikatsiooni saagise vahelist seost, on genereeritud rida mittepaarduvaid nukleotiide sisaldavaid, erineva NAS väärtusega praimereid.

Praimerite genereerimiseks viidi mittepaarduvad nukleotiidid sisse ainult pärisuunalise praimeri vastavatesse positsioonidesse. Vastassuunaline praimer jäeti antud töös muutmata.

Praimerite genereerimiseks on võetud aluseks Allergofood projekti raames eelnevalt disainitud ja katseliselt kontrollitud spetsiifilisusega makadaamiapähkli kloroplasti genoomi MatK geeni 199 bp pikkust regiooni amplifitseerivad praimerid (Mac0_F ja Mac0_R), mis on disainitud kasutades programmi *Primer3* (Untergasser *et al.*, 2012) ning NCBI Nucleotide andmebaasis olevat tervelehise makadaamia (*Macadamia integrifolia*, vili: makadaamiapähkel) geenijärjestust AY823204.1. Praimerite pikkused on pärisuunalisel praimeril 24 bp ja vastassuunalisel praimeril 22 bp. GC sisaldused (G ja C nukleotiide osakaalud järjestustes) on vastavalt 41,7% ja 45,5%, sulamistemperatuurid (T_m) 57,9°C ja 57,1°C.

Et kontrollida praimeri järjestuses oleva üksiku 3'-otsas paikneva mittepaardunud nukleotiidi mõju amplifikatsioonile meie katsetingimustes, on genereeritud pärisuunalised praimerid, mille erinevatesse 3'-otsa positsioonidesse on viidud erinevat tüüpi üksikud sihtmärkjärjestusega (makadaamiapähkli matK geeni järjestusega) mittepaarduvad nukleotiidid. Kuna üksiku mittepaardunud aluspaari mõju on oluline eelkõige polümeraasi seondumise ja praimerite pikendamise seisukohast, siis on antud töös viidud mittepaarduv nukleotiid 1.-6. positsiooni praimeri 3'-otsast. Et võrrelda erinevat tüüpi mittepaarduvate aluspaaride mõju amplifikatsioonile on genereeritud kõigi võimalike erinevat tüüpi mittepaarduvate nukleotiididega pärisuunaliste praimerite variante (võimalikke kombinatsioone on igas positsioonis kolm, näiteks kui praimeri ja sihtmärkjärjestuse nukleotiidide vahel on Watson-Crick paardumine A/T, siis võimalikud mittepaarduvate

nukleotiididega aluspaari variandid on G/T, C/T ja A/T). Võimalikud kombinatsioonid sõltusid sihtmärkjärjestusest (tabel 1).

Tabel 1. Ühe mittepaarduva nukleotiidiga pärisuunalised praimerid. Mittepaarduvad nukleotiidid paiknevad praimerid 3'-otsast 1. - 6. positsioonis. Mittepaarduva aluspaari tüübi G/T puhul on praimerid järjestuses nukleotiid G ning sihtmärkjärjestuse vastavas positsioonis T. Watson-Crick paardumise korral oleks antud positsioonis A/T paardumine. pu tähistab puriini ja py pürimidiini.

Praimeri ID	Pärisuunalise praimerid järjestus 5' → 3'	Mittepaarduva nukleotiidid positsioon	Mittepaarduva aluspaari tüüp	
Mc1GT_F	*****G	1	G/T	pu/py
Mc1TT_F	*****T	1	T/T	py/py
Mc1CT_F	*****C	1	C/T	py/py
Mc2TG_F	*****T*	2	T/G	py/pu
Mc2GG_F	*****G*	2	G/G	pu/pu
Mc2AG_F	*****A*	2	A/G	pu/pu
Mc3AA_F	*****A**	3	A/A	pu/pu
Mc3GA_F	*****G**	3	G/A	pu/pu
Mc3CA_F	*****C**	3	C/A	py/pu
Mc4GT_F	*****G***	4	G/T	pu/py
Mc4TT_F	*****T***	4	T/T	py/py
Mc4CT_F	*****C***	4	C/T	py/py
Mc5CA_F	*****C****	5	C/A	py/pu
Mc5GA_F	*****G****	5	G/A	pu/pu
Mc5AA_F	*****A****	5	A/A	pu/pu
Mc6AC_F	*****A*****	6	A/C	pu/py
Mc6CC_F	*****C*****	6	C/C	py/py
Mc6TC_F	*****T*****	6	T/C	py/py

Et uurida, millist mõju avaldavad amplifikatsioonile kaks erinevates positsioonides paiknevat mittepaardunud aluspaari praimerid ja sihtmärkjärjestuse vahel, on genereeritud pärisuunalised praimerid, millel on lisaks 3'-otsas (1. - 6. positsioonis) paiknevale tugeva mõjuga mittekomplementaarset aluspaari (puriin/puriin või pürimidiin/pürimidiin tüüpi) või nõrga mõjuga mittekomplementaarset aluspaari (puriin/pürimidiin või pürimidiin/puriin tüüpi) moodustavale mittepaarduvale nukleotiidile üks nõrga mõjuga mittekomplementaarset aluspaari moodustav mittepaarduv nukleotiid. Kuna kõiki võimalikke kahe mittepaardunud aluspaari asukoha kombinatsioone on väga palju, vaatleme vaid juhte, kus esimene mittepaarduv aluspaar paikneb praimerid 3'-otsa läheduses (1. - 6. positsioonis) ja teine 3., 6., 10. või 23. positsioonis 3'-otsast (tabel 2).

Tabel 2. Kahe mittepaarduva nukleotiidiga pärisuunalised praimerid. Esimene, praimer 3'-otsale lähemal paiknev, mittepaarduv nukleotiid paikneb praimerite 3'-otsast 1. - 6. positsioonis, teine mittepaarduv nukleotiid 3., 6., 10. või 23. positsioonis. pu tähistab puriini, py tähistab pürimidiini.

Praimeri ID	Pärisuunalise praimeri järjestus 5' → 3'	Mittepaarduva nukleotiidi positsioon	Mittepaarduva aluspaari tüüp	
Mc31NT_F	*****C*T	3 + 1	C/A + T/T	py/pu + py/py
Mc31NN_F	*****C*G	3 + 1	C/A + G/T	py/pu + pu/py
Mc61NT_F	*****A****T	6 + 1	A/C + T/T	pu/py + py/py
Mc61NN_F	*****A****G	6 + 1	A/C + G/T	pu/py + pu/py
Mc101NT_F	*****G*****T	10 + 1	G/T + T/T	pu/py + py/py
Mc101NN_F	*****G*****G	10 + 1	G/T + G/T	pu/py + pu/py
Mc231NT_F	*T*****T	23 + 1	T/G + T/T	py/pu + py/py
Mc231NN_F	*T*****G	23 + 1	T/G + G/T	py/pu + pu/py
Mc32NT_F	*****CA*	3 + 2	C/A + A/G	py/pu + pu/pu
Mc32NN_F	*****CT*	3 + 2	C/A + T/G	py/pu + py/pu
Mc62NT_F	*****A***A*	6 + 2	A/C + A/G	pu/py + pu/pu
Mc62NN_F	*****A***T*	6 + 2	A/C + T/G	pu/py + py/pu
Mc102NT_F	*****G*****A*	10 + 2	G/T + A/G	pu/py + pu/pu
Mc102NN_F	*****G*****T*	10 + 2	G/T + T/G	pu/py + py/pu
Mc232NT_F	*T*****A*	23 + 2	T/G + A/G	pu/py + pu/pu
Mc232NN_F	*T*****T*	23 + 2	T/G + T/G	pu/py + py/pu
Mc63NT_F	*****A**A**	6 + 3	A/C + A/A	pu/py + pu/pu
Mc63NN_F	*****A**C**	6 + 3	A/C + C/A	pu/py + py/pu
Mc103NT_F	*****G*****A**	10 + 3	G/T + A/A	pu/py + pu/pu
Mc103NN_F	*****G*****C**	10 + 3	G/T + C/A	pu/py + py/pu
Mc233NT_F	*T*****A**	23 + 3	T/G + A/A	pu/py + pu/pu
Mc233NN_F	*T*****C**	23 + 3	T/G + C/A	pu/py + py/pu
Mc64NT_F	*****A*T***	6 + 4	A/C + T/T	pu/py + py/py
Mc64NN_F	*****A*G***	6 + 4	A/C + G/T	pu/py + pu/py
Mc104NT_F	*****G*****T***	10 + 4	G/T + T/T	pu/py + py/py
Mc104NN_F	*****G*****G***	10 + 4	G/T + G/T	pu/py + pu/py
Mc234NT_F	*T*****T***	23 + 4	T/G + T/T	pu/py + py/py
Mc234NN_F	*T*****G***	23 + 4	T/G + G/T	pu/py + pu/py
Mc65NT_F	*****AA****	6 + 5	A/C + A/A	pu/py + pu/pu
Mc65NN_F	*****AC****	6 + 5	A/C + C/A	pu/py + py/pu
Mc105NT_F	*****G****A****	10 + 5	G/T + A/A	pu/py + pu/pu
Mc105NN_F	*****G****C****	10 + 5	G/T + C/A	pu/py + py/pu
Mc235NT_F	*T*****A****	23 + 5	T/G + A/A	py/pu + pu/pu
Mc235NN_F	*T*****C****	23 + 5	T/G + C/A	py/pu + py/pu
Mc106NT_F	*****G***C*****	10 + 6	G/T + C/C	pu/py + py/py
Mc106NN_F	*****G***A*****	10 + 6	G/T + A/C	pu/py + pu/py
Mc236NT_F	*T*****C*****	23 + 6	T/G + C/C	py/pu + py/py
Mc236NN_F	*T*****A*****	23 + 6	T/G + A/C	py/pu + pu/py
Mc2310NT_F	*T*****T*****	23 + 10	T/G + T/T	py/pu + py/py
Mc2310NN_F	*T*****G*****	23 + 10	T/G + G/T	py/pu + pu/py

Et kontrollida käesolevas töös leitud NAS väärtuste ja amplifikatsiooni saagise vahelist seost, on genereeritud erineva arvu ja 3'-otsast erineval kaugusega paiknevate, mittepaarduva nukleotiididega pärisuunalised praimerid, mille NAS skoori väärtused varieeruvad vahemikus 1 kuni 21 (lisa 1). Vahemik on valitud eeldades, et skoori 21 juures on mittepaardunud aluspaaride mõju nii praimeril ja järjestuse vahelise dupleksi stabiilsusele kui ka polümeraasi seondumise ja praimerite pikendamise efektiivusele juba piisavalt suur, et tuvastatavat produkti hulka (hinnatuna silma järgi geelelektroforeesi bändi intensiivsuse järgi) amplifikatsioonil ei tekiks. Sihtmärkjärjestustega täies ulatuses komplementaarse järjestusega praimeril (näiteks vastassuunalise praimeril) NAS väärtus on 0.

Erineva NAS väärtusega genereeritud praimerite järjestustes olevad mittepaardunud nukleotiidid tekitavad ainult nõrga mõjuga mittepaardunud aluspaare, kuna spetsiifilisuse seisukohast on just need kõige ohtlikumad (sama arvu ja positsiooniga tugeva mõjuga või sama arvu ja positsiooniga nõrga mõjuga aluspaaride esinemisel praimeril ja järjestuse vahel võib amplifikatsiooni tulemus olla väga erinev, kuid oht amplifikatsiooniks on suurem eelkõige nõrga mõjuga mittepaarduvate aluspaaride puhul).

Genereeritud praimerid on kontrollitud võimalike dimeeride ja sekundaarstruktuuride moodustamise osas veebiserveri DINAMelt (Markham ja Zuker, 2005) rakenduste *Quickfold* ja *Two state melting (hybridization)* abil, kasutades järgnevaid vaheväärtustest erinevaid parameetriväärtusi: soolakontsentratsioon $[Na^+]$ 10 mM ja $[Mg^{++}]$ 2,5 mM ning järjestuste kontsentratsiooni (*strand concentration*) 0,4 μM . Mittepaarduvate nukleotiididega praimerite genereerimisel on püütud vältida G ja C nukleotiidide sisalduse väärtuse liigset varieerumist praimerite vahel ning pärisuunalise praimeril pikkus on jäetud konstantseks (24 nukleotiidi), et vältida lisafaktorite mõju amplifikatsioonile.

Et kontrollida DNA ahelate hübridisatsioonitugevust iseloomustava Gibbsi vabaenergia muudu võimet ennustada amplifikatsiooni toimumist, leiti veebiserveri DINAMelt (Markham ja Zuker, 2005) rakenduse *Two state melting (hybridization)* abil kõikide genereeritud mittepaarduvate nukleotiididega pärisuunaliste praimerite sihtmärkjärjestusega seondumise Gibbsi vabaenergia muudu (ΔG_m) väärtus (kasutades parameetriväärtusi: soolakontsentratsioon $[Na^+]$ 10 mM ja $[Mg^{++}]$ 2,5 mM ning praimerite kontsentratsiooni (*strand concentration*) 0,4 μM) ning leiti selle erinevus ($\Delta \Delta G$) võrreldes täies pikkuses komplementaarse praimeril vabaenergia muuduga ($\Delta G_0 = -24,3$) ehk $\Delta \Delta G = \Delta G_m - \Delta G_0$ (lisa 1). Kuna vastassuunaline praimer on täies pikkuses komplementaarne sihtmärkjärjestusega, on selle $\Delta \Delta G$ väärtus 0.

2.2.2. DNA eraldus

Tervelehise makadaamia (*Macadamia integrifolia*, vili: makadaamiapähkel) DNA on eraldatud mahetoodete poest hangitud kooritud seemnest (pähklist). Pekani-hikkoripuu (*Carya illinoensis*, vili: pekanipähkel), kreeka pähklipuu (*Juglans regia*, vili: kreeka pähkel), hariliku sarapuu (*Corylus avellana*, vili: metspähkel/sarapuupähkel), hariliku maapähkli (*Arachis hypogaea*, vili: maapähkel, arahhis), arukase (*Betula pendula*), hariliku viinapuu (*Vitis vinifera*, vili: viinamari) DNA on eraldatud Tartu Ülikooli Botaanikaias, metsas, aias või aknalauul kasvanud taimedelt korjatud ja silikageeli abil kuivatatud taimede lehtedest.

Eraldamiseks kasutati CTAB meetodit (Doyle and Doyle, 1990). Eraldatud DNA lahustati 100 µl TE puhvris. DNA kvaliteeti ja kvantiteeti hinnati geelelektroforeesi abil 1% agarooosi sisaldusega geeli abil ning spektrofotomeetri NanoDrop® ND-1000 (NanoDrop Technologies; Wilmington, USA) abil vastavalt tootja juhendile. PCR reaktsiooniks tehti DNA lahjendused kuni kontsentratsioonini 20 ng/µl (katseteks genereeritud praimeritega) või 50 ng/µl (katseteks pekanipähkli spetsiifiliste praimeritega). DNA proovid säilitati -20 kraadi juures.

Eraldatud DNA proovide amplifitseeritavus on kontrollitud Allergofood projekti raames disainitud vastavate liikide spetsiifiliste PCR praimeritega.

2.2.3. Polümeraasi ahelreaktsiooni katsed

PCR reaktsioonides genereeritud praimeritega kasutati matriitsina makadaamiapähklist eelnevalt eraldatud DNA-d. Praimeritena kasutati reaktsioonis eelnevalt genereeritud sihtmärkjärjestusega mittepaarduvate nukleotiididega pärisuunalist praimerit ja sihtmärkjärjestuse seondumissaidiga täies pikkuses komplementaarse järjestustega vastassuunalist praimerit. Positiivse kontrollina kasutati makadaamia-spetsiifilisi, sihtmärkjärjestusega täies pikkuses komplementaarse järjestusega päri- ja vastassuunalist praimerit, negatiivne PCR kontroll sisaldas kõiki reaktsioonisegu komponente (sh makadaamia DNA), välja arvatud pärisuunalist praimerit.

PCR reaktsioonisegu mahuga 25 µl sisaldas 1x PCR puhvrit [Tris-HCl, (NH₄)₂SO₄; Solis BioDyne, Inc; Tartu, Eesti], 2.5 mM MgCl₂ (Solis BioDyne, Inc), 0.25 mM dNTP segu (Solis BioDyne, Inc), 0,4µM pärisuunalist praimerit (Microsynth AG); 0,4µM vastassuunalist praimerit (Microsynth AG). 5 U HOT FIREPol® DNA Polümeraasi (Solis BioDyne, Inc, Tartu, Eesti), 20 ng DNA-d ja MilliQ H₂O-d lõppmahuni.

Kõik PCR katsed viidi läbi võimalikult sarnastes tingimustes, et vältida PCR-i katsesüsteemist tulenevat varieeruvust. Kõikides geneeritud praimeritega teostatud reaktsioonides kasutati samu reagentide ja DNA lahjendusi. Tulemuste korratavuse kontrollimiseks tehti iga genereeritud praimeripaariga PCR reaktsioone vähemalt 3 korduses ja erinevatel päevadel.

Amplifitseerimine viidi läbi termotsükleris ABI GeneAmp PCRSystem 2700 (Applied Biosystems, Foster City, CA, USA) kasutades järgmist programmi: algne denaturatsioon 95°C juures 15 min; 35 tsüklit: denaturatsioon 95°C juures 30 sek, praimerite seondumine 54°C juures 30 sek, süntees (praimerite pikendamine, ekstensioon) 72°C juures 30 sek; süntees 72°C juures 10 min; säilitamine 4°C juures.

Amplifitseeritud fragmente kontrolliti geelelektroforeesil (1,3%-lisel agarosgeelil TBE puhvis (40 mM Tris-Cl, 1 mM EDTA, 45mM boorhape; pH 8.3), millele oli lisatud etiidiumbromiid). Elektroforees toimus pingel 115 V, 30 minutit. DNA visualiseeriti UV (ultraviolet) valguses.

Et anda poolkvantitatiivne hinnang iga erineva NAS skooriga praimeripaari amplifikatsiooni saagisele, anti geelelektroforeesi käigus tekkivate fragmendikogumite triibu (*band*) intensiivsusele silma järgi hinnang: 0 (Amp.puudub - amplifikatsioon puudub, geelelektroforeesipildi triip oli silmaga tuvastamatu), 1 (Amp.nõrk - nõrk amplifikatsioon, triip tuvastatav, kuid tunduvat nõrgem võrreldes positiivse kontrolliga) või 2 (Amp.tugev - tugev amplifikatsioon, positiivse kontrolliga umbes samaväärse intensiivsusega triip).

2.2.4. Arvutitarkvara arendus taksoni-spetsiifiliste praimerite disainiks

Käesolevas töös loodud programmid sihtmärkjärjestuse maskeerimiseks (`TargetConsensusMaker.py` ja `TargetMasker.py`) ning praimerite spetsiifilisuse hindamiseks ja sorteerimiseks (`PrimerSorter.py`) on kirjutatud programmeerimiskeeles Python, kasutades Pythoni versiooni 3.x ning on käivitavad Unix/Linux käsurealt. Programme on võimalik käivitada nii eraldiseisvate programmidenä (järjestuse maskeerimiseks või praimerite sorteerimiseks spetsiifilisuse alusel) kui ka üksteisele järgneva pideva töövoona (spetsiifiliste praimerite disainiks, kasutades lisaks programmi `Primer3` käsurea versiooni (Untergasser *et al.*, 2012)).

2.2.5. Taksoni-spetsiifiliste praimerite disaini metoodika rakendamine diagnostilise testi väljatöötamisel

Antud töös väljatöötatud praimeridisaini automatiseeritud töövoogu rakendati selleks, et disainida Allergofood projekti raames spetsiifilised praimerid pekanipähkli (pekani-hikkoripuu) tuvastamiseks toidust.

Praimerite disainiks otsustati kasutada pekanipähkli allergeeni 11S legumiini geeni järjestusi mRNA järjestusi, sest selle tuumagenoomi geenipiirkonna puhul on andmebaasis olemas piisavalt järjestusi paljude fülogeneetiliselt lähedaste ja pekani-hikkoripuu viljadega toidus sageli koos esinevate viljadega liikide (nt kreeka pähkclipuu, sarapuu) puhul ning järjestus varieerus piisavalt pekani-hikkoripuu ja kreeka pähkclipuu järjestuste vahel. Pekani-hikkoripuu mRNA järjestuste ekson-intron piirid ennustati kasutades Allergofood projekti raames sekveneeritud lähedase liigi sarapuu (*Corylus avellana*) genoomset järjestust, andmebaasis olevat sarapuu allergeeni Cor a 9 (11S globuliini valgu) geeni mRNA järjestusi ning NCBI BLAST programmi (blastn).

Praimerite disainil kasutatud pekani-hikkoripuu (*Carya illinoensis*) kui sihtmärkliigi 11S legumiini geeni järjestused (EU113051.1 ja EU113052.1) on võetud NCBI (*The National Center for Biotechnology Information*) andmebaasist Nucleotide. Söödavate viljadega fülogeneetiliselt lähedaste ja väga sarnaste DNA järjestustega mittesihhtmärkliikide kreeka pähkclipuu (*Juglans regia*), hariliku sarapuu (*Corylus avellana*) ja hariliku kastanipuu (*Castanea sativa*) järjestused (AY692446.1, AF449424.1, KF494372.1, JX036033.1, GQ925921.1, GQ925920.1) on leitud BLAST programmi (blastn) otsinguga NCBI nr/nt (*Nucleotide collection*) andmebaasi vastu (kasutades vaikeväärtusi). Mittesihhtmärkjärjestuste sarnasused sihtmärkjärjestusega (EU113051.1) olid kreeka pähkclipuu järjestusel 96%, hariliku sarapuu järjestustel 78% ja hariliku kastanipuu järjestustel 73%.

Kahe pekani-hikkoripuu järjestuse põhjal moodustati uus sihtmärkjärjestuste konsensusjärjestus, milles järjestuste vahel varieeruvad positsioonid olid maskeeritud tähega N, kasutades käesoleva töö raames loodud programmi TargetConsensusMaker.py.

Praimerite disainiks kasutati programmi Primer3 käsura versiooni (Untergasser *et al.*, 2012). Et vähendada mittespetsiifiliste praimerite kombinatsioone Primer3 väljundis, kasutati enne praimerite disaini käesolevas töös kirjutatud programmi/skripti TargetMasker.py (kasutades vaikeväärtusi), mille abil maskeeriti sihtmärkjärjestuses väikese tähega sihtmärk-taksoni (pekani-hikkoripuu) järjestuse ja mõne mittesihhtmärk-taksoni (kreeka pähkclipuu, hariliku sarapuu või hariliku kastanipuu) järjestuse vahel identsete 10 bp

pikkuste järjestuseosade otsanukleotiidid (1. ja 10. positsioonis). Programmi `Primer3` sisendfaili genereerimiseks kasutab `TargetMasker.py` eraldi tekstifaili, milles lisaks muude praimeridisaini parameetrite väärtustele sisaldub ka väikese tähega maskeerimist arvestavat parameeter (`PRIMER_LOWER_CASE_MASKING=1`) (näide lisas 5).

Programmi `Primer3` abil disainitud primereid sorteeriti ja seati spetsiifilisuse alusel pingeritta käesolevas töös loodud teise programmi `PrimerSorter.py` abil, kasutades programmi vaikeväärtusi (sisendfailitüüp – `Primer3` väljundfail, praimerite sorteerimine - vaid NAS skoori alusel).

Kõige spetsiifilisema praimeripaariga viidi läbi PCR katsed spetsiifilisuse kontrollimiseks lehtedest eraldatud DNA proovidega. PCR protokoll ja programm olid samad, mis PCR katsetes genereeritud mittepaarduvate nukleotiididega praimeritega (täpsem info on alalõigus 2.2.3). PCR matriitsina kasutati eelnevalt eraldatud pekani-hikkoripuu (*Carya illinoensis*, vili: pekanipähkel), kreeka pähklipuu (*Juglans regia*, vili: kreeka pähkel), hariliku sarapuu (*Corylus avellana*, vili: metspähkel ehk sarapuupähkel), tervelehise makadaamia (*Macadamia integrifolia*, vili: makadaamiapähkel), hariliku maapähkel (*Arachis hypogaea*, vili: maapähkel, arahhis), arukask (*Betula pendula*), hariliku viinapuu (*Vitis vinifera*, vili: viinamari) lehtedest või pähklist eraldatud DNA proove, kontsentratsiooniga 50 ng/μl.

2.3. Tulemused

2.3.1. Skoor praimerite pikendamise ennustamiseks

Senisest teaduskirjandusest on teada, et mida rohkem on mittepaardunud aluspaare ja mida rohkem paikneb neid 3'-otsa lähedal, seda nõrgem on praimer ja sihtmärk-DNA vahel moodustuv dupleks ja ebaefektiivsem on polümeraasi seondumine ning praimerite pikendamine. Samas ei taga praimerite 3'-otsas paiknev üksik mittepaardunud aluspaar veel amplifikatsiooni mittetoimumist.

Konkreetselt praimerite spetsiifilisust hindame oma töös muutuja NAS (*No Amplification Score*) väärtuse abil, mis peegeldab praimerite seondumise ja pikendamise efektiivsust konkreetse seondumiskoha puhul ning on leitav valemiga:

$$NAS = MM20 + 2 \times MM10 + 3 \times MM6 + 4 \times MM3,$$

kus MM20 on mittepaarduvate nukleotiidide arv praimerite 3'-otsast 11. - 20. positsioonis, MM10 on mittepaarduvate nukleotiidide arv 7. - 10. positsioonis, MM6 on mittepaarduvate nukleotiidide arv 4. - 6. positsioonis ning MM3 on mittepaarduvate nukleotiidide arv 1. - 3. positsioonis praimerite 3'-otsast.

NAS väärtus sõltub mittepaarduvate nukleotiidide koguarvust praimerite ja võimaliku seondumissaidi vahel ning annab suurema kaalu 3'-otsa lähedal olevatele mittepaarduvatele nukleotiididele, millel teadaolevalt on suurem mõju amplifikatsiooni toimumisele (võrreldes 5'-otsa mittepaarduvatele nukleotiididele), sest need mõjutavad lisaks praimerite seondumisele sihtmärk-DNA-ga ka polümeraasi seondumist dupleksiga ning praimerite pikendamist.

Praimeril, millel ei ole võrreldes konkreetse seondumissaidi järjestusega mitte ühtegi mittepaarduvat nukleotiidi ehk mis on terves pikkuses komplementaarne seondumissaidi järjestusega, on NAS väärtus 0. Mida suurem on väärtus, seda rohkem on üldiselt mittepaarduvaid nukleotiide ja seda rohkem praimerite 3'-otsa lähedal need paiknevad ning seda ebatõenäolisem on eeldatavalt praimerite seondumine ja pikendamine polümeraasi poolt. Maksimaalne võimalik skoori väärtus on 39 (olukord, kus praimerite ja võimaliku seondumissaidi vahelise dupleksi 20 esimest aluspaari, praimerite 3'-otsast alates, on mittepaarduvad).

Kuna mittepaardunud aluspaari tüübi mõju antud töös on uuritud vaid ühe ja kahe mittepaarduva nukleotiidiga praimerite puhul, piiratud arvu tüüpide kombinatsioonidega, ning teaduskirjanduses on tulemused erinevad ning kohati vastukäivad, siis ei ole me antud töös NAS leidmise valemisse integreerinud erinevate mittepaardunud aluspaaride tüüpide erinevat mõju amplifikatsioonile.

Kuna on näidatud, et 3'-otsa mittepaardunud olemasolu ei taga alati (eriti nõrgema mõjuga mittepaardunud alupaaride puhul) amplifikatsiooni mittetoimumist, siis ei ole NAS arvutamise valemis antud suuremat kaalu praimeris 3'-otsas 1. positsioonis paiknevale mittepaardunud nukleotiidile.

Antud töös on praimerite NAS väärtust kasutatud praimeripaaride spetsiifilisuse hindamiseks ja sorteerimiseks. Et arvuliselt hinnata praimeripaari spetsiifilisust tugine oma töös kummagi praimeris (pärisuunalise ja vastassuunalise praimeris) minimaalseimale leitud NAS väärtusele. NAS iseloomustab konkreetse praimeris ja konkreetse seondumissaidi vahelise seondumise ja praimerite pikendamise efektiivsust, seega kõige ohtlikum praimeris seondumissait etteantud mittesihhtmärkjärjestuste puhul on tõenäoliselt minimaalseima NAS väärtusega praimeris seondumiskoht.

Kuna ei ole veel täpselt teada, kuivõrd erinev on pärisuunalises ja vastassuunalises praimeris paiknevate sama arvu, positsiooni ja tüübiga mittepaarduvate nukleotiidide kombineeritud mõju amplifikatsioonile, siis on kummagi praimeris NAS väärtused leitud sama valemi alusel. Primeripaari spetsiifilisust iseloomustav NAS väärtus on antud töös arvutatud kui pärisuunalise ja vastassuunalise praimeris minimaalseima NAS väärtuse summa. Kui primeripaari mõlemad praimerid on täies pikkuses komplementaarsed mittesihhtmärkjärjestusega, on primeripaari summaarne NAS väärtus 0. Maksimaalne võimalik primeripaari summaarne NAS väärtus on 78 (olukorras, kus mõlema praimeris ja nende kõige tõenäolisemate mittespetsiifilise seondumissaitide vahelises dupleksis on 20 mittepaarduvat aluspaari (primeris 3'-otsas).

Kuna antud töös genereeritud praimerite puhul on mittepaarduvad nukleotiidid sisse viidud vaid pärisuunalise praimeris järjestusse, siis on vastassuunalise praimeris NAS väärtus 0 ning primeripaari summaarne NAS väärtus on võrdne pärisuunalise praimeris NAS väärtusega.

2.3.2. PCR katsete tulemused üksikuid mittepaarduvaid nukleotiide sisaldavate praimeritega

Amplifikatsiooni saagis praimeripaari puhul, mille pärisuunalise praimer järjekorras on üksik, 1. - 6. positsioonis (praimer 3'-otsast) paiknev, mistahes vaadeldud tüüpi mittepaardunud aluspaari moodustav mittepaarduv nukleotiid ning mille vastassuunaline praimer on täies pikkuses komplementaarne sihtmärkjärjestustega, oli peaaegu sama suur kui praimeripaari puhul, mille järjestustes mittepaarduvaid nukleotiide ei olnud (geelelektroforeesi pildid on lisas 2).

Ühe mittepaarduva nukleotiidiga pärisuunalise praimer variandid olid genereeritud makadaamiapähkli spetsiifilise pärisuunalise praimer alusel. Seetõttu olid erinevate moodustuvate mittepaarduvate aluspaaride kombinatsioonid pisut piiratud ja sõltusid sihtmärk-DNA järjestusest ning ei olnud võimalik kontrollida kõikides positsioonides olevate kõigi erinevat tüüpi üksikute mittepaarduvate aluspaaride kombinatsioonide mõju. Näiteks ei näinud me oma katsete põhjal praimer 3'-otsas paikneva puriin/puriin tüüpi mittepaardunud aluspaari mõju amplifikatsioonile. Kuna antud töös loodavates programmides on plaanis praimerite spetsiifilisust hinnata ainult mittesihhtmärkjärjestuste mittepaarduvate aluspaaride arvu ja positsiooni põhjal (arvestamata aluspaari tüüpi), siis on olulisem eelkõige teada nõrga mõjuga (puriin/pürimidiin või pürimidiin/puriin tüüpi) mittepaarduvate aluspaaride mõju amplifikatsioonile ning kontrollida kas üksik mittepaarduv aluspaar praimer 3'-otsa lähedal hoiab meie katsetingimustes ära (spetsiifilist või mittespetsiifilist) amplifikatsiooni. Antud töö tulemuste põhjal võivad ka praimerid, mille ühe praimer 3'-otsas on mittepaardunud nukleotiid (võrreldes mittesihhtmärk-järjestusega), anda (mittespetsiifilist) amplifikatsiooni.

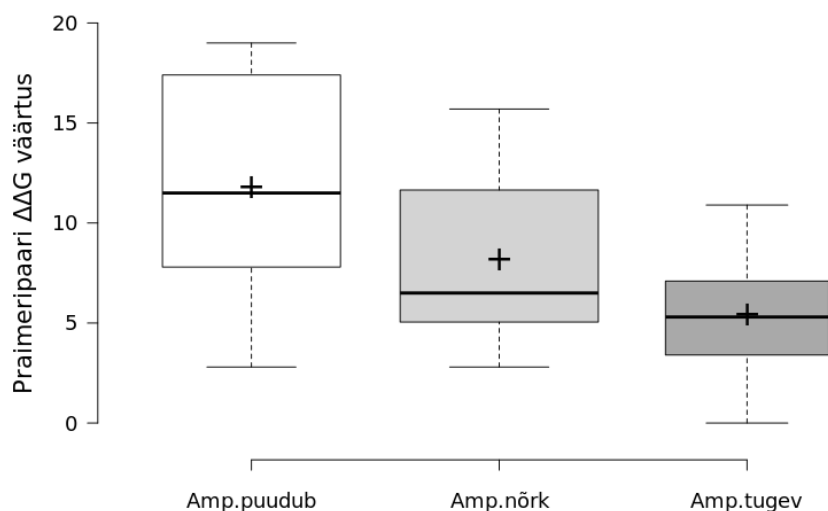
Antud töös saadud amplifikatsiooni tulemused praimeripaariga, mille pärisuunalises praimeris oli kaks mittepaarduvat nukleotiidi ning vastassuunaline praimer oli täies pikkuses komplementaarne sihtmärkjärjestusega, näitasid, et kaks mistahes positsioonides paiknevat mittepaarduvat aluspaari (millest 1 on nõrga mõjuga) üldiselt amplifikatsiooni ära ei hoi (vastavad geelelektroforeesi pildid on lisas 3). Meie katsetes ei toimunud amplifikatsiooni ainult olukorras, kus mittepaardunud nukleotiidid paiknesid 1. ja 3. positsioonis praimer 3'-otsast, kusjuures 3'-otsas paiknev mittepaardunud aluspaar oli T/T (pürimidiin/pürimidiin) tüüpi. Kaks nõrga mõjuga mittepaardunud aluspaari mistahes vaadeldud asukoha kombinatsioonides (sh 1. ja 3. positsioonis praimer 3'-otsast) amplifikatsiooni ära ei hoidnud. Samuti ei hoidnud amplifikatsiooni toimumist ära 1. positsioonis paiknev mittepaarduv

aluspaar T/T, juhul kui teine, nõrga mõjuga, mittepaarduv aluspaar paiknes praimeri 3'-otsast 6. positsioonis või kaugemal.

Antud töö tulemused näitavad, et kuni kaks mittepaarduvat aluspaari (sh 3'-otsas) praimeri ja selle seondumissaidi vahelises dupleksis, sõltumata asukohast, ei tarvitse amplifikatsiooni toimumist ära hoida.

2.3.3. Seos amplifikatsiooni saagise ja praimeripaari NAS väärtuse vahel

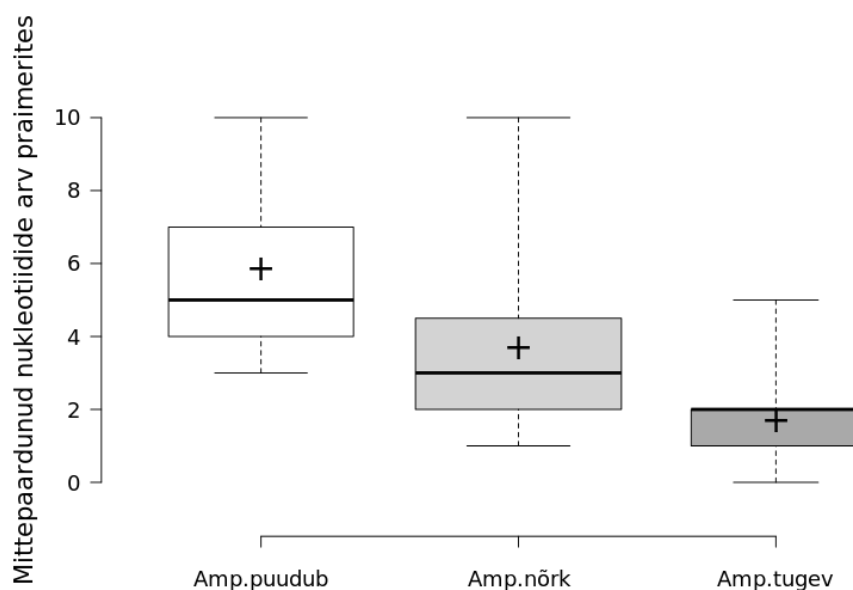
Praimeripaari amplifikatsiooni toimumise ennustamiseks kasutatakse sageli DNA ahelate hübridisatsiooni tugevust kirjeldavat Gibbsi vabaenergia muutu või järjestuste vaheliste mittepaarduvate nukleotiidide koguarvu. Joonis 3 kirjeldab meie katsetes erinevat PCR saagist (geelelektroforeesipildi triibu intensiivsust) näidanud praimeripaaride $\Delta\Delta G$ väärtuste varieerumist. $\Delta\Delta G$ väärtus on Gibbsi vabaenergia muudu erinevus mittepaarduvate nukleotiididega ja täies pikkuses komplementaarse praimeri vahel. Kuna sihtmärkjärjestusega komplementaarse vastassuunalise praimeri puhul on $\Delta\Delta G$ väärtus 0, siis praimeripaari summaarne $\Delta\Delta G$ väärtus võrdne pärisuunalise praimeri $\Delta\Delta G$ väärtusega.



Joonis 3. Praimerite $\Delta\Delta G$ väärtuste varieeruvus erineva saagisega amplifikatsiooni reaktsioonide puhul. $\Delta\Delta G$ väärtus on arvutatud kui Gibbsi vabaenergia muutuse erinevus sihtmärkjärjestusega täies pikkuses komplementaarse ja mittepaarduvate aluspaaridega seondumise vahel. Karp-vurrud (Box plot) diagrammi karpide keskmised jooned tähistavad $\Delta\Delta G$ väärtuste mediaani, kasti ülemine ja alumine joon vastavalt ülemist ja alumist kvartiili (75. ja 25. protsentiili), nn. „vurrud“ maksimaalset ja minimaalset väärtust. Rist tähistab väärtuste keskmist. Amplifikatsiooni saagist on hinnatud geoelektroforeesipildi triibu intensiivsuse alusel kolme tasemega: Amp.puudub – triip puudub või on tuvastamatu, Amp.nõrk – triip on oluliselt nõrgem kui positiivne kontroll (sihtmärkjärjestusega komplementaarse järjestusega praimeripaar), Amp.tugev – triibu intensiivsus on võrreldav positiivse kontrolli tulemusega.

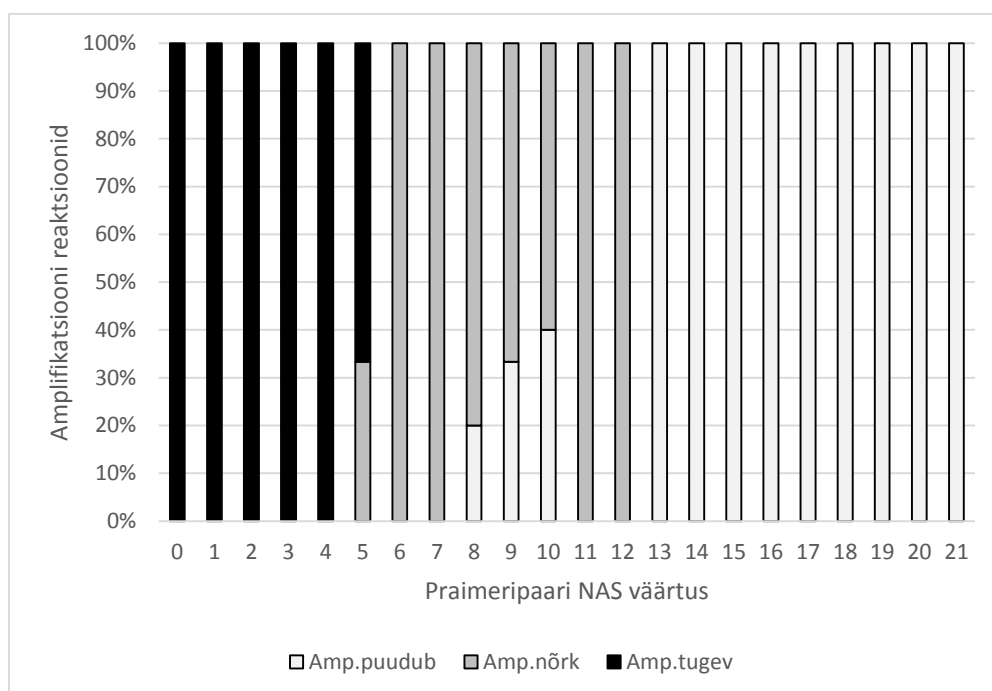
Kuigi keskmine $\Delta\Delta G$ väärtus erineb puuduva amplifikatsiooniga ja nõrga amplifikatsiooniga reaktsioonide puhul, ei saa konkreetse praimerit või praimeripaari $\Delta\Delta G$ väärtuse põhjal ennustada selle amplifikatsiooni toimumise tõenäosust või saagist (joonis 3).

Genereeritud pärisuunaliste praimerite järjestustes olevate mittepaarduvate nukleotiidide koguarvu varieerumist erineva saagisega amplifikatsioonireaktsioonide puhul on näha joonisel 4. Kuna vastassuunalise praimerit järjestuses mittepaarduvaid nukleotiide ei olnud, siis praimeripaari summaarne mittepaarduvate nukleotiidide koguarv oli võrdne mittepaarduvate nukleotiidide arvuga pärisuunalise praimerit järjestuses. Kuigi võiks arvata, et mida rohkem on praimerit ja seondumissaidi järjestuste vahel mittepaarduvaid aluspaare, seda ebatõenäolisem või väiksema saagisega on amplifikatsioon, siis antud töö tulemused näitasid, et ainult mittepaardunud aluspaaride arvu abil amplifikatsiooni toimumist ja saagist ennustada ei saa. Näiteks 3 - 10 mittepaarduva nukleotiidiga praimerit kasutades võib, sõltuvalt mittepaarduvate nukleotiidide asukohast, amplifikatsioon ebaõnnestuda, aga võib saada ka tuvastatava produktihulga (nõrga või koguni tugeva intensiivsusega triip) (joonis 4).



Joonis 4. Pärisuunalise praimerit järjestuses olevate mittepaarduvate nukleotiidide koguarvu varieeruvus erineva saagisega amplifikatsioonireaktsioonide puhul. Vastassuunaline praimerit oli täies pikkuses komplementaarne. Karp-vurrud (Box plot) diagrammi karpide keskmised jooned tähistavad mittepaarduvate nukleotiidide arvu väärtuste mediaani, kasti ülemine ja alumine joon vastavalt ülemist ja alumist kvartiili (75. ja 25. protsentiili), nn. „vurrud“ maksimaalset ja minimaalset väärtust. Rist tähistab väärtuste keskmist. Amplifikatsiooni saagist on hinnatud geelelektrofooresipildi triibu intensiivsuse alusel kolme tasemega: Amp.puudub – triip puudub või on tuvastamatu, Amp.nõrk – triip on oluliselt nõrgem kui positiivne kontroll (sihtmärkjärjestusega komplementaarse järjestusega praimeripaar), Amp.tugev – triibu intensiivsus on võrreldav positiivse kontrolli tulemusega.

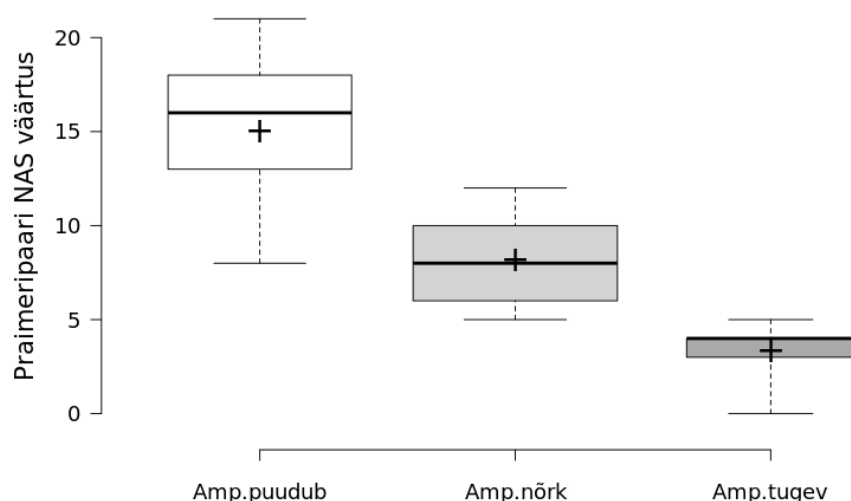
PCR katsete tulemused erineva summaarse NAS väärtusega praimeripaaridega näitasid, et mida suurem oli praimeripaari NAS väärtus, seda suurem oli ebaõnnestunud või väiksema saagisega amplifikatsioonireaktsioonide osakaal kogu reaktsioonide hulgast, mis konkreetse NAS väärtusega praimeritega tehti. Kui skoori väärtus oli väiksem kui 4, oli amplifikatsioon peaaegu sama efektiivne kui sihtmärkjärjestusega täies pikkuses komplementaarsete pärisuunalise praimeriga praimeripaari puhul. Kui praimeripaari skoori väärtus oli suurem kui 13, amplifikatsiooni ei toimunud (joonis 5). Erineva NAS väärtusega praimeripaaride PCR amplifikatsiooni geelipildid on toodud lisa 4.



Joonis 5. Erineva NAS väärtusega praimeripaaride erineva saagisega amplifikatsioonireaktsioonide osakaalud kõikidest konkreetse skooriga praimeritega tehtud reaktsioonidest. Amplifikatsiooni saagist on hinnatud geelelektroforeesipildi triibu intensiivsuse alusel 3 tasemega: Amp.puudub – triip puudub või on tuvastamatu, Amp.nõrk – triip nõrk või oluliselt nõrgem kui positiivse kontrolli (sihtmärkjärjestusega komplementaarne praimeripaar) puhul, Amp.tugev – triibu tugevus võrreldav positiivse kontrolli tulemusega.

Saadud tulemuste põhjal ei saa me usaldusväärselt määrata küll skoori väärtuse täpset lävendväärtust, millest alates amplifikatsiooni enam ei toimu, kuid näeme, et NAS väärtus on pöördvõrdelises seoses amplifikatsiooni toimumise tõenäosuse ja saagisega: mida suurem on praimeripaari summaarne NAS väärtus, seda väiksem on amplifikatsiooni saagis ja tõenäolisem amplifikatsiooni mittetoimumine.

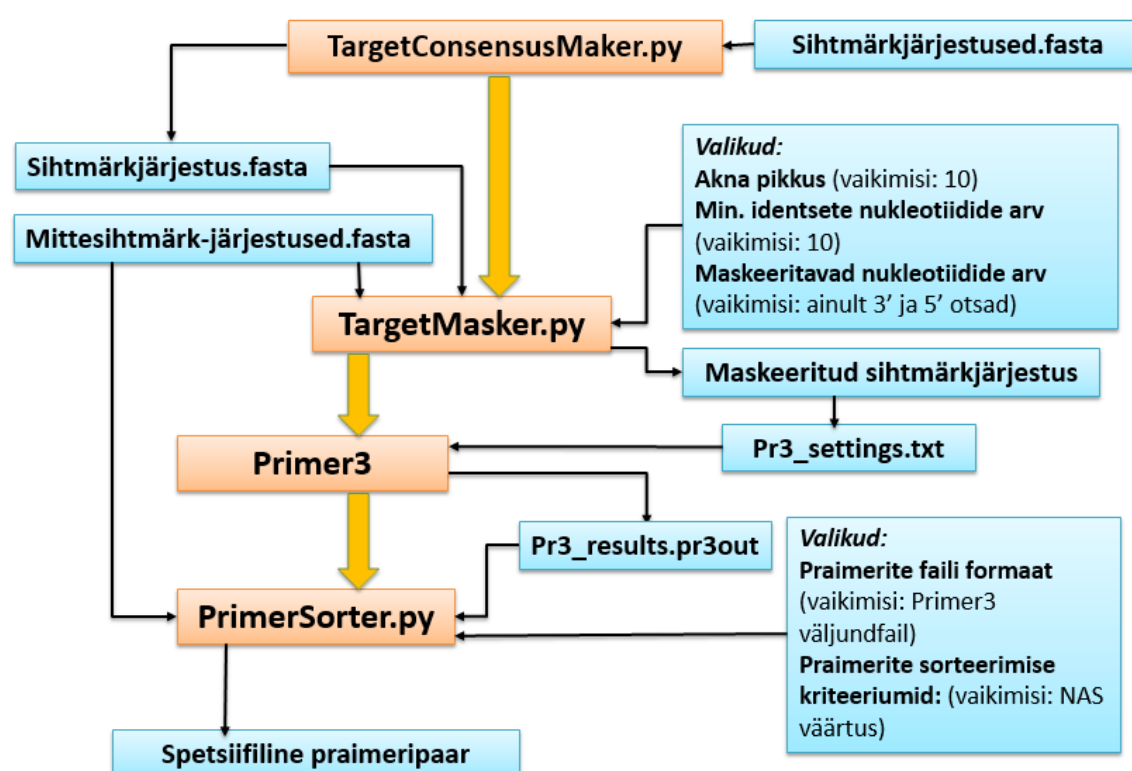
Joonis 6 kirjeldab detailsemalt erinevat PCR saagist (geelelektroforeesipildi triibu intensiivsust) näidanud praimerite NAS skoori väärtuste varieerumist. Geelelektroforeesi abil silmaga eristatav triip tuvastati skoori väärtuse juures väiksem kui 13. Samas oli NAS väärtuse juures 8 kuni 13 üksikuid reaktsioone, mille puhul silmaga tuvastatavat triipu ei tekkinud, kas silma järgi triibu intensiivsuse hindamise piiratusest või amplifikatsiooni toimumist mõjutavate muude lisategurite tõttu.



Joonis 6. Praimerite NA skoori väärtuste varieeruvus erineva amplifikatsiooni saagisega reaktsioonide puhul. Karp-vurrud (Box plot) keskmised jooned tähistavad NA skoori väärtuste mediaani, kasti ülemine ja alumine joon vastavalt ülemist ja alumist kvartiili (75. ja 25. protsentiili), nn. „vurrud“ maksimaalset ja minimaalset väärtust. Rist tähistab väärtuste keskmist. Amplifikatsiooni saagist on hinnatud geelelektroforeesi triibu intensiivsuse alusel kolme tasemega: Amp.puudub – triip ei ole tuvastatav, Amp.nõrk – triip on oluliselt nõrgem kui positiivne kontroll (sihtmärkjärjestusega komplementaarsete praimeritega), Amp.tugev – triibu intensiivsus on võrreldav positiivse kontrolli tulemusega.

2.3.4. Taksoni-spetsiifiliste PCR praimerite disaini meetoodika ja tarkvara

Spetsiifiliste praimerite disainiks on käesolevas töös loodud automatiseeritud neljast programmist koosnev töövoog (joonis 7). Programmid `TargetConsensusMaker.py`, `TargetMasker.py` ja `PrimerSorter.py` on väljatöötatud käesoleva töö käigus ja on kasutatavad nii eraldiseisvate programmidenä kui ka järjestikuse töövoona Linux/Unix keskkonna käsurealt ning eeldavad programmi Python olemasolu. Järjestikuse töövoona rakendamisel on vajalik praimerite disainiks ka programmi `Primer3` käsurea versiooni (Untergasser *et al.*, 2012) olemasolu.



Joonis 7. Spetsiifiliste praimerite disaini töövoog. Programmid on tähistatud roosa kastiga, sisend- ja väljundinfo sinise kastiga, programmide käivitamise järjekord oranžide nooltega. `TargetConsensusMaker.py` kasutab sisendina sihtmärgjärjestusi FASTA formaadis failina ja genereerib sihtmärgjärjestuste konsensusjärjestuse FASTA formaadis failina. `TargetMasker.py` kasutab sisendina sihtmärgjärjestuse ja mittesihthärgjärjestuste FASTA formaadis faile ja soovi korral ka parameeterväärtused (libiseva otsinguakna pikkus; minimaalsete mittepaarduvate nukleotiidide arv aknas, mille puhul maskeeritakse; maskeeritavate nukleotiidide arv sihtmärgjärjestuses) ning genereerib `Primer3` sisendfaili. `Primer3` abil disainitakse etteantud optimaalsete parameeterväärtustega praimeripaarid. Programm `PrimerSorter.py` kasutab sisendina praimerijärjestusi (vaikimisi `Primer3` väljundfailina) ja mittesihthärgjärjestusi FASTA formaadis faile ning väljastab praimeripaarid spetsiifilisuse alusel pingereas.

Praimerite disainiks sobiva genoomipiirkonna valimine, mittesihtmärk-taksonite määratlemine ja järjestuste hankimine andmebaasidest on olulised etapid taksoni-spetsiifiliste praimerite disainil. Kuna disainiks kasutatava geeniregiooni valimine nõuab iga taksoni-spetsiifilise testi puhul individuaalset lähenemist, sõltuvalt tuvastatavast taksonist, mittesihtmärk-taksonitest, andmebaasis olemasolevatest geenijärjestustest, siis ei ole seda automatiseeritud töövoogu integreeritud ning eeldatakse, et kasutajal on konkreetse taksoni-spetsiifilise testi seisukohast olulised sihtmärk- ja mittesihtmärkjärjestused FASTA formaadis failidena genoomiandmebaasidest juba hangitud.

Mida pikem on praimeridisainiks kasutatav sihtmärkjärjestus, seda suurem on tavaliselt praimeridisaini programmi poolt väljastatav võimalike praimeripaaride kombinatsioonide hulk. Teatud määral on võimalik seda arvu vähendada piirates väljundina näidatavat praimeripaaride arvu, piirates parameetrite väärtusi, keelates praimerijärjestuste kattuvust, kuid see võib omakorda filtreerida välja võimalikke väga häid kandidaate spetsiifiliste praimeritena. Et pigem välja filtreerida praimereid, mis võiksid anda mittespetsiifilist amplifikatsiooni mittesihtmärkliikidelt, on mõistlik enne praimeridisaini maskeerida praimeridisainiks ebasobivad regioonid sihtmärkjärjestuses (nt spetsiifiliste praimerite disaini puhul mittespetsiifilised praimeri 3'-otsa kohad). Antud töös on maskeerimiseks kirjutatud programm `TargetMasker.py`, mis võtab sisendina sisse sihtmärkjärjestuse faili (FASTA formaadis) ning mittesihtmärkjärjestuste faili (FASTA formaadis) ja otsib etteantud pikkusega libiseva akna abil järjestuse piirkondi (vaikimisi 10 nukleotiidi pikkuseid), mis on liiga sarnased sihtmärkjärjestuse ja mõne mittesihtmärkjärjestuse vahel ning leidmise korral (kasutaja poolt on määratud minimaalne lubatud identsete nukleotiidide arv aknas, mille puhul on järjestus liiga sarnane; vaikeväärtus on 10) maskeerib etteantud arvu nukleotiidide leitud järjestuseosa mõlemas otsas (vaikimisi ainult 3'- ja 5'-otsas paikneva ühe nukleotiidi) väikese tähega. `Primer3` puhul on võimalik vastava parameetri abil keelata disainida praimereid, mille 3'-ots paikneb väikese tähega tähistatud positsioonis. Programm `TargetMasker.py` genereerib väljundina programmi `Primer3` sisendfaili (kasutades selleks kasutaja poolt määratud disainiparameetri väärtusi sisaldavat alusfaili `Pr3settings_form.txt`), mis sisaldab praimerite disainiks vajalike parameetrite väärtusi ning sihtmärkjärjestust.

`TargetMasker.py` vaikeväärtusi (otsitava akna pikkus: 10, minimaalne lubatud identsete nukleotiidide arv aknas maskeerimiseks: 10, maskeeritavate nukleotiidide arv akna järjestuse 3'- ja 5'-otsast ühele või teisele poole: 0 – ainult 3'- ja 5'-ots) on võimalik muuta,

kuid tuleb arvestada, et see mõjutab maskeerimise rangust ning hiljem `Primer3` väljundina saadavate praimerite hulka. Maskeerimise optimaalne rangus on vajalik, et tagada siiski praimeridisainil piisav hulk praimeripaare, mille seast valida kõige spetsiifilisem kandidaat.

Kui sihtmärk-taksoni järjestusi on rohkem kui üks, tuleb enne `TargetMasker.py` käivitamist moodustada sihtmärkjärjestuste konsensusjärjestus kasutades näiteks käesoleva töö raames kirjutatud programmi `TargetConsensusMaker.py`, mis joondab järjestused kasutades programmi `ClustalW` (Larkin *et al.*, 2007) ja väljastab sihtmärkjärjestuste konsensusjärjestuse (FASTA formaadis failina), milles järjestuste vahel varieeruvate nukleotiididega positsioonides on nukleotiid tähistatud tähega N.

Praimerite disainil programmi `Primer3` (Untergasser *et al.*, 2012) abil (võib kasutada teisi praimeridisaini programme) leitakse konkreetse sihtmärk-taksoni järjestuse amplifitseerimiseks sobivad praimeripaarid. `Primer3` poolt väljastatud sobivad praimerid on kontrollitud dimeeride ja sekundaarstruktuuride moodustamise osas ning on reastatud selle alusel, kuivõrd kasutaja poolt määratud optimumväärtuste lähedased on parameetrite väärtused konkreetse praimeripaari puhul.

Et hinnata disainitud praimeripaaride spetsiifilisust ja sobivust konkreetse taksonspetsiifilise testi jaoks ning valida välja kõige sobivam praimeripaar enne PCR reaktsiooni läbiviimist laboris, on käesoleva töö raames kirjutatud eraldiseisev arvutiprogramm `PrimerSorter.py`. Programm tugineb praimeripaari spetsiifilisuse hindamisel käesolevas töös väljatöötatud skoorile NAS.

Programm võtab sisendina sisse mittesihtmärk-DNA järjestused (FASTA formaadis failina) ja praimerite järjestused (vaikimisi `Primer3` väljundfailina) ning soovi korral ka sorteeritud praimereid sisaldava uue faili nime. Programm otsib nii pärisuunalise kui vastassuunalise praimeri kõige ohtlikumat (minimaalsema NAS väärtusega) seondumist üle kõigi mittesihtmärkjärjestuste ning hindab konkreetse taksoni-spetsiifilise testi jaoks kõige sobivamaks praimeripaari, mille kahe praimeri puhul leitud minimaalsete NAS väärtuste summa on maksimaalne ehk praimeripaari, mille kahe praimeri kõige ohtlikumate seondumiste pikendamise efektiivsused oleksid kõige madalamad. Kuna pärisuunaline praimer võib käituda ka vastassuunalise praimerina ja vastupidi, otsib programm praimerijärjestusega nii komplementaarseid kui pöördkomplementaarseid seondumiskohti. Programm eeldab, et praimerite pikkused on vähemalt 10 nukleotiidi.

NAS väärtustel põhinev `PrimerSorter.py` programm töötab eeldusel, et mittesihtmärkjärjestuste failis on kõik järjestused sellised, mis on konkreetse taksoni-

spetsiifilise testi seisukohast olulised mittesihtmärkjärjestused (st millelt mitte üheltki järjestuselt praimerid amplifikatsiooni anda ei tohiks).

`PrimerSorter.py` võimaldab valikute abil praimerijärjestusi ette anda ka teatud kujul esitatud praimerijärjestusi sisaldava tekstifailina (praimeripaarid eraldi ridadel, igal real Tab-eraldatud pärisuunalise ja vastassuunalise praimeri järjestus), näiteks juhul, kui praimerid ei ole disainitud `Primer3` abil). Praimerite sorteerimisel spetsiifilisuse alusel ning pingerea koostamisel on võimalik arvestada/nõuda lisaks NAS väärtusele ka praimerite 3'-otsa mittepaarduva nukleotiidi olemasolu (praimeri ja mittesihtmärkjärjestuse vahel).

Kuna programmid `TargetMasker.py` ja `PrimerSorter.py` ei tugine järjestuste võrdlemisel järjestuste joondusele, siis ei ole ohtu, et joondamisvead mõjutavad saadavat tulemust. Kuna programmid ei tugine BLAST algoritmile, suudavad need väga täpselt määrata praimeri ja mittesihtmärkjärjestuse vaheliste mittepaardunud nukleotiidide arvu libisevas aknas ning ei oma BLAST otsingu piirangut minimaalse mittepaardunud aluspaaride vahelise kauguse osas.

Programm `PrimerSorter.py` väljastab praimeripaarid sorteerituna praimeripaari summaarse NAS skoori väärtuse alusel. Väljundina saadavas tekstiformaadis on iga praimeripaari puhul toodud kummagi praimeri järjestused, võimalikud ohtlikuimad (minimaalseima NAS väärtusega) seondumissaidid (järjestused, nende genoomiandmebaasi ID, liigi nimi), kummagi praimeri NA skoori väärtused ja 3'-otsas paikneva mittepaarduva nukleotiidi olemasolu või puudumine (vastavalt väärtusega 1 või 0).

Maskeerimiseks või sorteerimiseks kuluv aeg programmidel sõltub järjestuste pikkusest, arvust ning leitavatest identsete järjestuseosade arvust. Ühe sihtmärkjärjestuse ja umbes 60 mittesihtmärkjärjestuse puhul pikkusega umbes 220 bp kulus maskeerimiseks umbes paar minutit, praimerite disainiks ja sorteerimiseks paar sekundit.

Efektiivsete koodi ridade arv (ei sisalda kommentaariridu ja tühje ridu) on programmil `TargetConsensusMaker.py` 57, programmil `TargetMasker.py` 122 ja programmil `PrimerSorter.py` 596.

2.3.5. Väljatöötud tarkvara ja metoodika abil disainitud spetsiifilised praimerid pekanipähkli tuvastamiseks

Pekanipähklit tuvastavate spetsiifiliste praimerite disainimiseks kasutati järjestikuse töövoona programme: `TargetConsensusMaker.py`, `TargetMasker.py`, `Primer3` ja `PrimerSorter.py`. Antud töövoos programmide mugavamaks järjekorras käivitamiseks olid programmid `TargetConsensusMaker.py`, `TargetMasker.py`, `PrimerSorter.py` ja `Primer3` ning programmi `TargetMasker.py` käivitamiseks vajalikud failid failid (nt `Pr3settings_form.txt`, FASTA formaadis sihtmärkjärjestuste ja FASTA formaadis mittesihtmärkjärjestuste failid) paigutatud samasse kataloogi.

Programmi `TargetConsensusMaker.py` abil moodustati sihtmärkjärjestuste konsensusjärjestus, milles oli N tähega maskeeritud positsioone u. 1%.

Programmi `TargetMasker.py` väljundina saadi `Primer3` sisendfail, mis sisaldas nii kasutaja poolt määratud praimeridisaini parameetrite väärtusi (lisa 5) kui ka maskeeritud sihtmärkjärjestust. Kuna kreeka päklikipuu ja pekani-hikkoripuu vastavate geenide järjestused erinevad vaid üksikute nukleotiide poolest, jäi `TargetMasker.py` töö tulemusena jäi pekani-hikkoripuu järjestusse maskeerimata positsioone, kuhu oli võimalik praimerite disainil paigutada praimerite 3'-otsi, pärast maskeerimist järele 22.

`Primer3` abil disainiti kõik võimalikud kasutaja poolt määratud, biokeemiliste ja muude parameetriväärtuste poolest sobivate ja dimeeride ning sekundaarstruktuuride moodustumis osas kontrollitud praimeripaaride kombinatsioonid. Praimeridisaini parameetrite väärtused olid valitud selliselt, et need oleksid kooskõlas hiljem laboris kasutatava PCR protokolliga ja programmiga ning väljundina saadavad praimerid tuvastaksid võimalikult suure tundlikkuse ja spetsiifilisusega pekanipähkli DNA-d töödeldud toiduainetest. Pekanipähkli (pekani-hikkoripuu) spetsiifiliste praimerite disainil saadi kokku 10 võimalikku praimeripaari (saadavate praimeripaaride hulk sõltub järjestuse pikkusest, lubatavatest parameetriväärtustest, maskeerimise rangusest).

Kasutades praimeridisainiks maskeerimata järjestust, saadi väljundina 515 paari praimereid. Sõltuvalt järjestusest võib see arv ületada ka kasutaja poolt määratud maksimaalset väljundina lubavat praimeripaaride arvu, mistõttu võivad spetsiifilised praimeripaari kandidaadid jääda ka väljundfailist välja. Samuti erinesid pärisuunaliste ja vastassuunaliste praimerite kombinatsioonid praimeripaarides, sõltuvalt sellest, kas kasutati maskeerimata või

maskeeritud järjestust. Kui võrreldi spetsiifiliste praimerite leidmise tulemust eelnevalt maskeeritud ja maskeerimata sihtmärkjärjestust kasutades, selgus, et praimeridisainil saadavate praimeripaaride pärisuunaliste ja vastassuunaliste praimerite kombinatsioonid olid erinevad (isegi kui lubada praimerite kattuvust paarides). Maskeeritud järjestust kasutades kõige spetsiifilisemateks osutunud praimeripaaride pärisuunalised ja vastassuunalised praimerid ei esinenud maskeerimata järjestuse kasutamisel saadud praimeripaaride kombinatsioonides praimeripaarina koos (kumbki praimer oli eraldi teistes kombinatsioonides) ning maskeerimata järjestuse puhul leitud kõige spetsiifilisemate praimeripaaride summaarne NAS väärtus oli 10. See näitab mittespetsiifiliste praimerite vähendamise vajalikkust eelneva maskeerimisetapi abil.

Programmi `PrimerSorter.py` töö tulemusena osutusid kolmeks kõige spetsiifilisemaks praimeripaarid summaarsete NAS väärtustega vastavalt 13 (mõlemal praimeril oli 3'-otsas mittepaarduv nukleotiid), 13 (ühel praimeril 3'-otsa mittepaarduv nukleotiid) ja 11 (ühel praimeril 3'-otsa mittepaarduv nukleotiid). Samad praimerid olid `Primer3` väljundfaili pingereas vastavalt 8., 6. ja 9. positsioonis (kokku oli 10 praimeripaari).

Joonisel 8 on toodud `PrimerSorter.py` abil leitud kõige spetsiifilisema praimeripaari (suurima praimerite summaarse NAS väärtusega paari) pärisuunalise ja vastassuunalise praimerite mittepaarduvate nukleotiidide paiknemine minimaalseima NAS väärtusega seondumiskoha (*Juglans regia*) puhul.

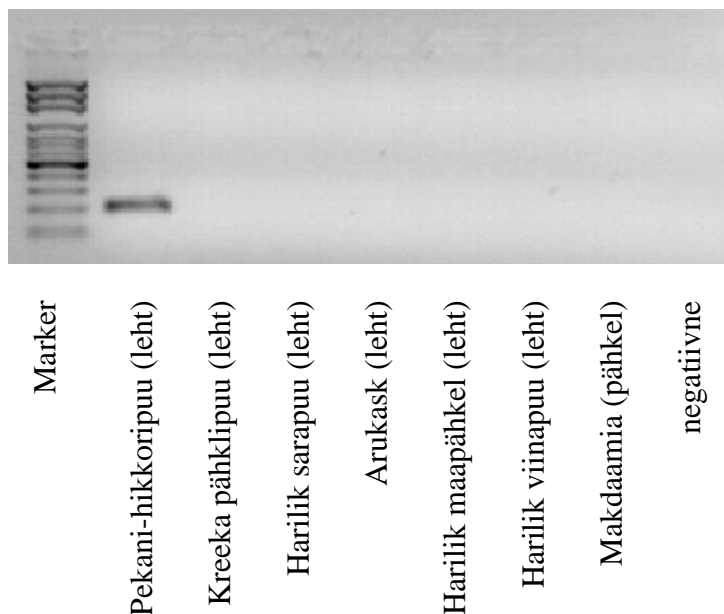
Pärisuunaline praimer (<i>C. illinoensis</i>) 5' → 3'		*****
<i>Juglans regia</i> :		*****N
Vastassuunaline praimer (<i>C. illinoensis</i>) 5' → 3'		*****
<i>Juglans regia</i> :		*****N***N**N

$NAS_{sum} = NAS_F + NAS_R = 4 + 9 = 13$
--

Joonis 8. Pekanipähkli (*Carya illinoensis*) spetsiifiliste praimerite ja minimaalseima NAS väärtusega seondumiskohtade (*Juglans regia*) järjestuste vaheliste mittepaarduvate nukleotiidide (N) paiknemine. NAS_F on pärisuunalise praimerite minimaalne NAS skoor, NAS_R on vastassuunalise praimerite minimaalne NAS skoor, NAS_{sum} on praimeripaari summaarne NAS skoor.

Kõige spetsiifilisema praimeripaari spetsiifilisust kontrolliti ka Primer-BLAST (Ye *et al.*, 2012) otsinguga kogu NCBI Nr andmebaasi vastu ning ei leitud mitte ühtegi mittespetsiifilisi seondumist (seondumise mittekuvamiseks pidi praimerid ja järjestuse vahel olema vähemalt 6 mittepaardunud aluspaari, millest 3 pidi paiknema praimerid 3'-otsast 1. - 3. positsioonis).

Viidi läbi ka spetsiifilisust kontrollivad PCR katsed eelnevalt eraldatud DNA proovidega (kreeka pähklipuu (*Juglans regia*, vili: kreeka pähkel), hariliku sarapuu (*Corylus avellana*, vili: metspähkel/sarapuupähkel), maapähkli (*Arachis hypogaea*, vili: maapähkel, arahhis), arukase (*Betula pendula*), hariliku viinapuu (*Vitis vinifera*, vili: viinamari) ning tervelehine makadaamia (*Macadamia integrifolia*), positiivne kontroll pekani-hikkoripuu (*Carya illinoensis*, vili: pekanipähkel)). Saadud tulemused näitavad, et praimerid olid pekanipähkli spetsiifilised. Amplifikatsioonil tekkis oodatava pikkusega produkt pekani-hikkoripuu puhul, teiste liikide (sh väga sarnase DNA järjestusega kreeka pähklipuu) puhul produkti ei tekkinud (joonis 9).



Joonis 9. Amplifikatsioon pekani-hikkoripuu spetsiifiliste praimeritega. Amplifikatsioonil tekkis produkt pekanipähkli puhul. Teiste liikide puhul produkti ei tekkinud.

Edasised laboratoorsed katsed on vajalikud, et kontrollida veel praimerite tundlikkust ja spetsiifilisust DNA-de segu ja reaalseste töödeldud, pekanipähklit sisaldavate ja mittesisaldavate, toiduaineproovide puhul.

2.4. Arutelu

Mittespetsiifiline amplifikatsioon sõltub kõige enam erinevustest (mittepaardunud aluspaaridest) praimeri ja mittesihtmärkjärjestuse vahel. Liigispetsiifiliste praimerite disainil on tihti olukord, kus sihtmärkjärjestus erineb mittesihtmärkjärjestustest ainult üksikute nukleotiidide poolest. Sellisel juhul on oluline teada mitme, mis positsioonis ja mis tüüpi mittepaardunud alupaari esinemisel praimeri ja mittesihtmärkjärjestuse vahel võib veel toimuda mittespetsiifiline amplifikatsioon. Üldiselt on teada, et mida rohkem on mittepaardunud aluspaare praimeri ja järjestuse vahel ja mida rohkem paiknevad need praimeri 3'-otsa lähedal, seda suurem on nende mõju nii hübridisatsioonile kui praimerite pikendamisele ning seeläbi kogu amplifikatsioonile (tulemuseks on väiksem efektiivsus ja saagis) (Lefever *et al.*, 2013; Wright *et al.*, 2014), kuid üksikute, ühe-kahe erinevat tüüpi mittepaardunud aluspaari mõju kohta on erinevates katsetes saadud erinevaid, kohati vastukäivaid tulemusi (Day *et al.*, 1999; Ayyadevara *et al.*, 2000; Wu *et al.*, 2009; Stadhouders *et al.*, 2010; Wright *et al.*, 2014; Lefever *et al.*, 2013).

Olemasolevaid teadmisi mittepaarduvate aluspaaride mõjust praimerite seondumise ja pikendamise efektiivsusele rakendati praimerite seondumise ja pikendamise efektiivsust hindava skoori väljatöötamisel. Kuna kõiki erinevat tüüpi mittepaarduvate aluspaaride kombinatsioone me oma töös ei uurinud ja seni avaldatud andmed erinevat tüüpi mittepaarduvate aluspaaride mõju kohta amplifikatsioonile on vastukäivad, tugineb antud töös välja töötatud praimerite seondumise ja pikendamise efektiivsust hindav skoor (NAS) ainult mittepaardunud aluspaaride arvul ja positsioonil. Kuna mittepaardunud aluspaarid 3'-otsa lähedal mõjutavad samaaegselt nii praimerite ja sihtmärkjärjestuste hübridisatsiooni kui ka praimerite pikendamist polümeraasi poolt, annab NAS suurema kaalu 3'-otsas paiknevatele mittepaardunud aluspaaridele. Mida väiksem on praimeri NAS väärtus, seda suurem on eeldatavalt praimeri seondumise ja selle pikendamise efektiivsus konkreetse seondumissaidi puhul. Kuna ei ole täpselt teada, kas ja millisel määral on erinev mittepaarduvate aluspaaride mõju pärisuunalise või vastassuunalise praimeri seondumisele ja pikendamisele, on kummagi praimeri NAS väärtus arvutatud samadel alustel. Praimeripaari NAS väärtus on kummagi praimeri NAS väärtuse summa.

Et kontrollida erinevas positsioonis olevate mittepaarduvate aluspaaride mõju amplifikatsioonile meie katsetingimustes, tehti käesoleva töö raames mõned PCR katsed erinevas positsioonis paiknevate üksikuid mittepaarduvaid nukleotiide sisaldavate praimeritega. Kuna sarnase DNA järjestusega liikide eristamisel on eelkõige oht, et järjestuste

vahel on liiga vähe ja nõrga mõjuga mittepaarduvad aluspaarid, siis keskenduti antud töös eelkõige nõrga mõjuga aluspaaride kombinatsioonide mõju uurimisele. Seni avaldatud teadusartiklites on leitud, et üksik mittepaardunud aluspaar praimeris 3'-otsas ei hoia alati ära amplifikatsiooni ning selle mõju avaldub eelkõige praimeris pikendamise efektiivsuse languses ja sõltub mittepaardunud aluspaari tüübist (Lefever *et al.*, 2013; Li *et al.*, 2004; Huang *et al.*, 1992; Day *et al.*, 1999) ja katsetingimustest. Meie katsete tulemused näitasid, et ühe praimeris 3'-otsa lähedal (1. - 6. positsioonis) paikneva üksiku mittepaarduva nukleotiidiga praimeripaari puhul toimub amplifikatsioon mistahes vaadeldud erinevat tüüpi aluspaari puhul. Saadud tulemused olid kooskõlas varasemate töödega, milles on leitud, et üksik G/T, C/T või T/T tüüpi mittepaardunud aluspaar praimeris 3'-otsas avaldab üldiselt nõrka mõju praimerite pikendamise efektiivsusele (Day *et al.*, 1999; Kwok *et al.*, 1990; Li *et al.*, 2004; Stadhouders *et al.*, 2010). Kuigi mitmetes varasemates katsetes on leitud, et T/T (pürimidiin/pürimidiin) 3'-otsa mittepaardunud aluspaari esinemisel praimeris järjestuses on amplifikatsioon oluliselt nõrgem kui täies pikkuses komplementaarsete järjestustega praimerite puhul (Huang *et al.*, 1992; Ayyadevara *et al.*, 2000), siis meie katsetes oli amplifikatsioon mõlemal juhul peaaegu sama tugev, mis on kooskõlas mitmete teiste varasemate tööde tulemustega (Kwok *et al.*, 1990; Day *et al.*, 1999; Li *et al.*, 2004; Stadhouders *et al.*, 2010). Kuna antud töös vaadeldud mittepaardunud aluspaari tüüpide kombinatsioonide hulk oli piiratud ning sõltus kasutatud sihtmärkjärjestusest, siis ei saanud me katsega uurida teadaolevalt tugeva mõjuga mittepaardunud 3'-otsa aluspaaride (näiteks A/G, G/A, C/C) mõju amplifikatsioonile.

Uurides oma töös kahe erinevas positsioonis paikneva mittepaardunud aluspaari mõju amplifikatsioonile, leidsime et, kui lisaks 3'-otsas (1. - 6. positsioonis) paiknevale mittepaarduvale nukleotiidile praimeris järjestuses on teine nõrga mõjuga (pürimidiin/puriin või puriin/pürimidiin tüüpi) mittepaarduv nukleotiid (3., 6., 10 või 23. positsioonis), siis üldjuhul amplifikatsioon toimub. Amplifikatsioonil tuvastatavat saagist ei tekkinud ainult juhul, kui mõlemad mittepaardunud aluspaarid paiknesid praimeris 3'-otsast esimese kolme positsiooni hulgas, kusjuures üks neist paiknes 3'-otsas ja oli T/T (pürimidiin/pürimidiin) tüüpi, mis kirjanduse andmetel on tugevama mõjuga mittepaarduv aluspaar (Lefever *et al.*, 2013). Varem leitud tulemused kahe mittepaarduva aluspaari mõju kohta on olnud väga erinevad, alates sellest, et kahe mittepaardunud aluspaari puhul amplifikatsiooni ei toimu (Stadhouders *et al.*, 2010) kuni selleni, et teatud tüüpi mittepaardunud aluspaaride puhul amplifikatsioon toimub (Wright *et al.*, 2014). Wright *et al.*, (2014) ja Lefever *et al.*, (2013) leidsid, et kahe mittepaardunud aluspaari paiknemisel esimese kolme positsiooni hulgas praimeris 3'-otsast, kusjuures üks neist paikneb 3'-otsas, on amplifikatsioon oluliselt pidurdatud. Meie katsetes

kahe mittepaarduva nukleotiidiga praimeritega toimus oluline amplifikatsiooni inhibitsioon vaid T/T tüüpi 3'-otsa mittepaarduva aluspaari puhul ning olukorras, kus mõlemad mittepaarduvad aluspaarid olid nõrga mõjuga (pürimidiin/puriin ja puriin/pürimidiin tüüpi), toimus amplifikatsioon peaaegu sama suure lõppsaagisega kui täies pikkuses komplementaarsete praimerite puhul. Käesolevas töös saadud tulemuste põhjal võib järeldada, et üks-kaks mittepaarduvat aluspaari (sh 3'-otsas paiknevat) praimerit ja DNA järjestuse vahel ei tarvitse amplifikatsiooni ära hoida.

Antud töös leitud praimerite pikendamise efektiivsust kirjeldavat skoori NAS kasutati taksoni-spetsiifilise testi jaoks disainitud praimeripaaride spetsiifilisuse hindamiseks. Praimeri kõige ohtlikum mittespetsiifiline seondumine on kõige väiksema NAS väärtusega seondumine mittesihhtmärkjärjestustel ning kõige spetsiifilisemal praimeripaaril on pärisuunalise ja vastassuunalise praimerit väikseimate NAS skoori väärtuste summa kõige suurem.

Praimeripaari summaarse NAS skoori ja amplifikatsiooni saagise vahelise seose kontrollimiseks tehtud PCR katsete tulemused näitasid, et praimeripaari summaarse NAS skoori ja amplifikatsiooni saagise vahel on pöördvõrdeline seos: mida suurem on paari summaarne NAS skoor, seda väiksema saagisega toimub amplifikatsioon ning seda ebatõenäolisem on amplifikatsiooni toimumine (joonised 3 ja 4). Praimeripaari mittepaarduvate aluspaaride koguarvu ja amplifikatsiooni saagise või dupleksi stabiilsust iseloomustava $\Delta\Delta G$ ja amplifikatsiooni saagise vahel nii selget seost me oma töös tehtud katsete tulemuste puhul ei näinud (joonis 5).

Kuigi tundub ahvatlev väita, et alates praimeripaari NAS väärtusest 13 amplifikatsiooni ei toimu, ei saa saadud tulemuste põhjal veel välja tuua konkreetset praimeripaari NAS lävendväärtust, mille juures (mittespetsiifiline) amplifikatsioon toimub või enam ei toimu. Kuna amplifikatsiooni saagise hindamine tugines silma järgi geelelektroforeesi pildi triibu intensiivsuse hindamisel (3 diskreetset taset), siis võis juhtuda, et amplifikatsioon küll toimus, kuid geelelektroforeesi triip oli silmaga taustast eristamatu (näiteks NAS väärtuste 8 – 12 puhul). Samuti ei tarvitse tuvastatava või mittetuvastatava amplifikatsiooni saagisega NAS lävendväärtus olla tugeva mõjuga mittepaarduvate aluspaaridega praimerite puhul (millega antud töös katseid ei tehtud) sama, mis nõrga mõjuga mittepaarduvate aluspaaride korral ning tõenäoliselt muutub tugemate mittepaarduvate aluspaaride puhul amplifikatsiooni saagis tuvastamatuks madalama NAS väärtuse juures. Kuna lähedaste liikide eristamisel on eelkõige oht, et disainitud praimerite ja mittesihhtmärkjärjestuste vahelisi erinevusi on liiga vähe ja nõrga mõjuga, siis oli mõistlik kontrollida väljatöötatud skoori väärtuste ja amplifikatsiooni saagise seost katseliselt eelkõige nõrga mõjuga mittepaarduvaid nukleotiide sisaldavate praimeritega.

Kuigi leitud skoori ei saa kasutada selleks, et öelda, kas konkreetne praimeripaar annab või ei anna mittespetsiifilist amplifikatsiooni mittesihtmärkjärjestustelt, saime väljatöötatud skoori edukalt kasutada praimerite mittespetsiifilise seondumise ja pikendamise efektiivsuse hindamiseks, praimeripaaride hindamiseks, sorteerimiseks ning pingeritta seadmiseks spetsiifilisuse alusel ning vastava tarkvara väljatöötamisel.

Et välja töötatud skoor suudaks arvestada ka erinevat tüüpi mittepaarduvate aluspaaride erinevat mõju praimerite seondumisele ja pikendamisele, oleks vaja teha veel lisakatseid erinevat tüüpi (eriti tugeva mõjuga) mittepaarduvaid aluspaare moodustavate praimeritega. Samuti on võimalik NAS väärtuse täpsemaks arvutamiseks uurida sama arvu ja positsiooniga mittepaarduvate nukleotiidide mõju erineva pikkusega praimerite puhul ning sama arvu ja positsiooniga mittepaardunud nukleotiidide erinevat mõju esinemisel pärisuunalises või vastassuunalises praimeris või jaotatuna mõlemas praimeris vahel.

Seni ei olnud loodud sobivat metoodikat ja tarkvara taksoni-spetsiifilise testi jaoks praimerite disainiks ning eriti praimeripaaride sorteerimiseks ja valimiseks taksoni-spetsiifilise testi jaoks. Seetõttu oli oluline käesolevas töös töötada välja skoor, mis aitaks hinnata praimerite spetsiifilisust nende (mittespetsiifilise) seondumise ja pikendamise efektiivsuse kaudu. Antud töös loodud metoodikale tuginev ja erinevatest Linux/Unix käsurea programmidest koosneva automatiseeritud töövoos erinevad etapid aitavad tagada, disainida ja valida konkreetse PCR-il põhineva taksoni-spetsiifilise testi jaoks kõige sobivamat praimeripaari.

Kuna paljude eukariootsete organismirühmade (nt taimede) takson-spetsiifiliste praimerite disainil ei saa veel kasutada täisgenoomijärjestusi (nende vähesuse tõttu andmebaasides), siis ei ole disainil piiravaks teguriks mitte programmide kiirus, vaid eelkõige nende täpsus praimerite spetsiifilisuse hindamisel. Kuna `PrimerSorter.py` ei ütle kasutajale, kas konkreetse praimeris puhul mittespetsiifiline seondumine ja amplifikatsioon toimub või ei toimu, siis ei pea programm teadma praimerite seondumise või amplifikatsiooni (mitte)toimumise lävendväärtust NAS skoori, mittepaardunud aluspaaride arvu või nende positsiooni kohta ning seetõttu ei nõua kasutajalt vastavate parameetrite sisestamist nagu mõned olemasolevad programmid (nt `Primer-BLAST`, Ye *et al.*, 2012). Käesolevas töös loodud programmid `PrimerSorter.py` ja `TargetMasker.py` ei kasuta sarnaste järjestuseosade otsimiseks järjestuste joondust ega tugine BLAST algoritmile (Altschul *et al.*, 1997) ning seetõttu ei ole otsingu tulemus mõjutatud joonduse vigadest ega BLAST algoritmi piirangutest nagu mitmete olemasolevate programmide puhul (nt `Primer-BLAST`, Ye *et al.*, 2012; `QuantPrime`, Arvidsson *et al.*, 2008; `PRIMEGENS`, Xu *et al.*, 2002; `Primique`, Fredslund

ja Lange, 2007; SP-Designer, Villard ja Malausa, 2013). Samuti ei ole, erinevalt mitmetest teistest programmidest, käesolevas töös loodud programmide puhul mittesihhtmärkjärjestuste valik piiratud programmi enda poolt ette antud genoomijärjestustega (nt MFEprimer-2.0, Qu et al., 2012), vaid kasutaja saab mittesihhtmärkjärjestused FASTA formaadis failina ise kokku panna.

Käesolevas töös loodud metoodikat ja tarkvara rakendati praimerite disainil taksoni-spetsiifilise testi välja töötamiseks allergiat põhjustavate viljadega taimeliigi pekani-hikkoripuu (*Carya illinoensis*, vili: pekanipähkel ehk pekaanipähkel) tuvastamiseks toidust (projekti Allergofood raames). Genoomiandmebaasis NCBI olevad liigi pekani-hikkoripuu geenide järjestused on väga sarnased sama sugukonna liigi kreeka pähklipuu (*Juglans regia*, vili: kreeka pähkel) geenide järjestustega ning tihti on keeruline leida genoomijärjestuste andmebaasides olevate piiratud geenijärjestuste arvu ja sarnasuse tõttu kõige sobivamat kohta neid liike eristavate praimerite disainiks. PCR katsed disainitud praimeritega näitasid, et väljatöötatud programmide abil disainitud pramerid on pekanipähkli spetsiifilised ning loodud automatiseeritud töövoogu saab kasutada edukalt liigispetsiifiliste praimerite disainiks ning kõige spetsiifilisema kandidaatpaari väljaalimiseks.

Väljatöötatud metoodikat on võimalik kasutada tõenäoliselt mistahes taksoni-spetsiifilise (nt liigi-, perekonna-, sugukonna-spetsiifilise) testi jaoks PCR praimerite disainiks mistahes rakenduste tarbeks (näiteks eukariootsete allergeensete, toksiliste või ohustatud liikide DNA tuvastamiseks toiduainetest ja keskkonnaproovidest) PCR meetodi abil.

Kokkuvõte

Käesoleva töö raames loodi metoodika ja tarkvara PCR praimerite disainiks ja välja valimiseks taksoni-spetsiifiliste diagnostiliste testide jaoks.

Et hinnata praimerite seondumise ja pikendamise efektiivust, loodi skoor NAS (*No Amplification Score*), mis tugines varem avaldatud teaduskirjanduse andmetel praimeride ja seondumissaidi vahelise dupleksi mittepaarduvate aluspaaride arvu ja asukoha kohta. Väljatöötatud skoori rakendamiseks praimerite taksoni-spetsiifilisuse hindamisel tehti rida eksperimentaalseid katseid:

- 1) PCR katsete abil üksikuid mittepaarduvaid nukleotiidide sisaldavate praimeritega selgitati erinevas positsioonis paiknevate mittepaarduvate aluspaaride mõju amplifikatsioonile meie katsetingimustes ning leiti, et üks või kaks mittepaarduvat aluspaari praimeride ja seondumissaidi vahelise dupleksi mistahes positsioonides (sh praimeride 3'-otsas) ei tarvitse ära hoida amplifikatsiooni;
- 2) praimeripaari NAS skooride ja amplifikatsiooni saagise vahelise seose uurimiseks tehtud PCR katsed erinevate skoori väärtustega praimeripaaridega näitasid, et praimeripaari NAS skoori ja amplifikatsiooni saagise vahel on pöördvõrdeline seos;

Taksoni-spetsiifiliste praimerite disainiks, sorteerimiseks ja välja valimiseks loodi käesolevas töös neljast programmist koosnev automatiseeritud töövoog, mis hõlmab endas kolme käesolevas töös kirjutatud programmi (`TargetConsensusMaker.py` – sihtmärkjärjestuste vahelise varieeruvuse maskeerimiseks, `TargetMasker.py` – sihtmärk- ja mittesihhtmärkjärjestuste vaheliste väga sarnaste regioonide maskeerimiseks sihtmärkjärjestuses, `PrimerSorter.py` – disainitud praimeripaaride spetsiifilisuse hindamiseks ja sorteerimiseks NAS väärtuse alusel) ning kasutab lisaks sobivate praimeripaaride genereerimiseks ning sekundaarsutruktuuride ja dimeeride võimaliku moodustumise kontrollimiseks programmi `Primer3` käsurea versiooni (Untergasser *et al.*, 2012).

Väljatöötatud metoodika ja tarkvara rakendamisel allergeensete viljadega taimeliigi (pekani-hickoripuu) spetsiifiliste praimerite disainil õnnestus disainida spetsiifilised praimerid diagnostilise testi jaoks pekanipähkli tuvastamiseks toidust. Väljatöötatud metoodikat ja tarkvara on edaspidi võimalik kasutada mistahes taksoni-spetsiifilise testi jaoks vajaminevate praimerite disainiks või eraldiseisvaid programme kasutada sihtmärkjärjestuse maskeerimiseks, disainitud praimerite ohtlikumate mittespetsiifiliste seondumiskohtade leidmiseks mittesihhtmärkjärjestustelt ja praimeripaaride spetsiifilisuse hindamiseks ning valimiseks.

Development of methodology for taxon-specific primer design and its application for detection of DNA of allergenic plant

Kairi Raime

Summary

One of the most reliable and cost-effective ways to detect an organism is to carry out a polymerase chain reaction (PCR). Taxon-specific diagnostic tests use a pair of taxon-specific PCR primers to selectively amplify target sequence from the DNA sample. The biological DNA sample usually contains DNA sequences of various, sometimes phylogenetically very close, target and nontarget taxons. The specificity of amplification is achieved by the specificity of the two PCR primers, which mainly depends on the number of mismatches and the locations of mismatches of nonspecific binding sites.

Although there are several programs available for designing and checking the specificity of PCR primers, most of them use the sequence alignment tools, based on homology search, for finding nonspecific binding sites and do not consider the important factors (mismatches of 3' end of primer sequence) which influence the extension efficiency of the PCR primer and the specificity of primer.

Since there was not available a good methodology or software for designing or software for checking the specificity of primers for taxon-specific PCR based diagnostic test, we developed these in this work.

To assess the extension efficiencies of the primer, the score NAS (*No Amplification Score*) was created. NAS is based on the number of mismatches and the locations of the mismatches (between a primer and its potential binding site). Checking the correlation between NAS value of primerpair and amplification yield, we confirmed that there is inverse relationship between NAS value and amplification yield. NAS was applied to assess the taxon-specificity of the designed primer pairs. The most probable nonspecific binding site of the primer has minimal NAS value. The taxon-specificity of primerpair was described with the sum of minimal NAS value of both forward and reverse primer. We also developed the software for checking the specificity of primers, sorting and ranking designed primer pairs by NAS value.

For designing and ranking of primer pairs for taxon-specific diagnostic test the automated work pipeline, consisting of four programs (three created in this work) was developed: `TargetConsensusMaker.py` – for making consensus sequence of target

sequences, `TargetMasker.py` – for masking nonspecific positions of target sequence, `Primer3` (Untergasser *et al.*, 2012) - for generating primer pairs with optimal values of predefined parameters and checking of primers secondary structures formation, `PrimerSorter.py` – for sorting and ranking the primer pairs according to the NAS value.

We applied developed methodology and software for designing primer pairs for detecting allergenic plant species *Carya illinoensis* (pecan nut) with allergenic seeds. As the results we received taxon-specific primers for detection of pecan nut in food. Developed methodology and software can be used in the future for PCR primer design for taxon-specific diagnostic tests for various purposes. The developed programs can be used also independently for target sequence masking, for testing nonspecific bindings of primers or for ranking primer pairs.

Tänuavaldused

Käesoleva magistritöö raames läbiviidud uuringuid rahastati Haridus- ja Teadusministeeriumi grantidest IUT34-11 ja SF0180026s09 ning Euroopa Liidu Euroopa Regionaalarengu Fondi programmide "Genoomika tippkeskus" ja „Allergofood“.

Soovin tänada oma juhendajaid professor Maido Remmi ja Triinu Kõressaart heade nõuannete ja igakülgse toetuse eest magistritöö kirjutamisel.

Suured tänud Liina Nagirnajale Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudist PCR-alaste teadmiste, oskuste ja kogemuste jagamise eest ning Ülle Aarnale ja Tatjana Ojale Tartu Ülikooli Ökoloogia- ja Maateaduste Instituudist nõuannete ja juhendamise eest DNA eraldamisel taimedest.

Suur tänu Tartu Ülikooli Botaanikaia abivalmis töötajatele ning Tarmo Puurandile enda hoolega kasvatatud taimede lehtede loovutamise eest DNA eraldamiseks.

Kirjanduse loetelu

- Allawi, H.T. and SantaLucia, Jr. J. (1997). Thermodynamics and NMR of internal GT mismatches in DNA. *Biochemistry* **36**: 10581-10594.
- Allawi, H.T. and SantaLucia, Jr. J. (1998a). Nearest-neighbor thermodynamics of internal AC mismatches in DNA: sequence dependence and pH effects. *Biochemistry* **37**: 9435-9444.
- Allawi, H.T. and SantaLucia, Jr. J. (1998b). Nearest-neighbor thermodynamic parameters for internal GA mismatches in DNA. *Biochemistry* **37**: 2170-2179.
- Allawi, H.T. and SantaLucia, Jr. J. (1998c). Thermodynamics of internal CT mismatches in DNA. *Nucleic Acids Res* **26**: 2694-2701.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**(17): 3389–3402.
- Andreson, R., Reppo, E., Kaplinski, L. and Remm, M. (2006). GENOMEMASKER package for designing unique genomic PCR primers. *BMC Bioinformatics* **7**: 172.
- Aranyi, T., Varadi, A., Simon, I. and Tusnady, G.E. (2006). The BiSearch web server. *BMC Bioinformatics* **7**: 431.
- Arezi, B., Xing, W., Sorge, J.A. and Hogrefe, H.H. (2003). Amplification efficiency of thermostable DNA polymerases. *Anal Biochem* **321**: 226–235.
- Arvidsson, S., Kwasniewski, M., Riano-Pachon, D.M. and Mueller-Roeber, B. (2008). QuantPrime - a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics* **9**: 465
- Ayyadevara, S., Thaden, J.J. and Shmookler Reis, R.J. (2000). Discrimination of primer 3'-nucleotide mismatch by Taq DNA polymerase during polymerase chain reaction. *Anal Biochem* **284**: 11–18.
- Boom, N., DeWindt, W., Verstraete, W. and Top, E.M. (2002). Evaluation of nested PCR-DGGE (denaturing gradient gel electrophoresis) with group-specific 16S rRNA primers for the analysis of bacterial communities from different wastewater treatment plants. *FEMS Microbiology Ecology* **39**: 101-112.
- Boyle, B., Dallaire, N. and MacKay, J. (2009). Evaluation of the impact of single nucleotide polymorphisms and primer mismatches on quantitative PCR. *BMC Biotechnol* **9**: 75.
- Breslauer, K.J., Frank, R., Blöcker, H. and Marky, L.A. (1986). Predicting DNA duplex stability from the base sequence. *Proceedings of the National Academy of Sciences* **83**: 3746–3750.
- Bryksin, A.V. and Matsumura, I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques* **48**(6): 463–465.

- Catalano, C.E., Allen, D.J. and Benkovic, S.J. (1990). Interaction of Escherichia coli DNA polymerase I with azidoDNA and fluorescent probes: identification of protein–DNA contacts. *Biochemistry* **29**: 3612–3621.
- Cha, R.S., Zarbl, H., Keohavong, P., and Thilly, W.G. (1992). Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. *PCR Methods Appl* **2**: 14–20.
- Cha, R.S. and Thilly, W.G., (1995) Specificity, efficiency, and fidelity of PCR, in: C.W. Dieffenbach, G.S. Dveksler (Eds.), *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,
- Christopherson, C., Sninsky, J. and Kwok, S. (1997). The effects of internal primer-template mismatches on RT-PCR: HIV-1 model studies. *Nucleic Acids Res* **25**: 654–658.
- Chuang, L.Y., Cheng, Y.H. and Yang, C.H. (2013). Specific primer design for the polymerase chain reaction. *Biotechnol Lett* **35**(10): 1541-1549.
- Croning, M.D., Fricker, D.G., Komiyama, N.H. and Grant, S.G. (2010). Automated design of genomic Southern blot probes. *BMC Genomics* **11**: 74
- Day, J.P., Bergstrom, D., Hammer, R.P. and Barany, F. (1999). Nucleotide analogs facilitate base conversion with 3' mismatch primers. *Nucleic Acids Res* **27**: 1810–1818.
- Detera, S.D., Becerra, S.P., Swack, J.A. and Wilson, S.H. (1981). Studies on the mechanism of DNA polymerase alpha. Nascent chain elongation, steady state kinetics, and the initiation phase of DNA synthesis. *J Biol Chem* **256**(13): 6933-6943.
- Doyle, J.J. and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–15.
- Duitama, J., Kumar, D.M., Hemphill, E., Khan, M., Mandoiu, I.I. and Nelson, C.E. (2009). PrimerHunter: a primer design tool for PCR-based virus subtype identification. *Nucleic Acids Res* **37**(8): 2483–2492.
- Elnifro, E.M., Ashshi, A.M., Cooper, R.J. and Klapper P.E. (2000). Multiplex PCR: optimization and application in diagnostic virology. *Clin Microbiol Rev* **13**(4): 559–570.
- Ferrie, R.M., Schwarz, M.J., Robertson, N.H., Vaudin, S., Super, M., Malone, G. and Little, S. (1992). Development, multiplexing, and application of ARMS tests for common mutations in the CFTR gene. *Am J Hum Genet* **51**(2): 251–262.
- Fisher, P.A. and Korn, D. (1981). Ordered sequential mechanism of substrate recognition and binding by KB cell DNA polymerase α . *Biochemistry* **20**: 4560-4569.
- Fredslund, J. and Lange, M. (2007). Primique: automatic design of specific PCR primers for each sequence in a family. *BMC Bioinformatics* **8**: 369.
- Fu, Q., Ruegger, P., Bent, E., Chrobak, M., and Bomeman, J. (2008). PRISE (PRImer SElector): software for designing sequence- selective PCR primers. *J Microbiol Methods* **72**: 263–267.
- Gray, D.M. (1997). Derivation of nearest-neighbor properties from data on nucleic acid oligomers. II. Thermodynamic parameters of DNA.RNA hybrids and DNA duplexes. *Biopolymers* **42**(7): 795–810.

- Haas, S., Vingron, M., Poustka, A., Wiemann, S. (1998). Primer design for large scale sequencing. *Nucleic Acids Res* **26**(12): 3006–3012
- Higuchi, R., Fockler, C., Dollinger, G. and Watson, R. (1993). Kinetic PCR analysis: real-time monitoring of DNA amplification reactions, *Biotechnology* **11**: 1026–1030.
- Huang, M.M., Arnheim, N. and Goodman, M.F. (1992). Extension of Base mispairs by Taq dna-polymerase – implications for single nucleotide discrimination in PCR. *Nucleic Acids Res* **20**: 4567–4573.
- Huang, Y.T., Yang, J.I., Chrobak, M. and Borneman, J. (2014). PRISE2: software for designing sequence-selective PCR primers and probes. *BMC Bioinformatics* **15**: 317.
- Jarman, S.N., Gales, N.J., Tierney, M., Gill, P.C. and Elliott, N.G. (2002). A DNA-based method for identification of krill species and its application to analysing the diet of marine vertebrate predators. *Molecular Ecology* **11**: 2679-2690.
- Jarman, S.N., Deagle, B.E. and Gales, N.J. (2003). Group-specific PCR for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology* **13**(5): 1313-1322.
- Johnson, S.J. and Beese, L.S. (2004). Structures of mismatch replication errors observed in a DNA polymerase. *Cell* **116**:13.
- Kool, E.T. (1997). Preorganization of DNA: Design Principles for Improving Nucleic Acid Recognition by Synthetic Oligonucleotides. *Chem Rev* **97**(5): 1473-1488.
- Kool, E.T. (2001). Hydrogen bonding, base stacking, and steric effects in dna replication. *Annu Rev Biophys Biomol Struct* **30**: 1-22.
- Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C. and Sninsky, J.J. (1990). Effects of primer template mismatches on the polymerase chain-reaction – human-immunodeficiency-virus type-1 model studies. *Nucleic Acids Res* **18**: 999–1005.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.
- Lefever, S., Pattyn, F., Hellemans, J. and Vandesompele, J. (2013). Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clin Chem* **59**: 1470–1480.
- Lexa, M. and Valle, G. (2003). PRIMEX: rapid identification of oligonucleotide matches in whole genomes. *Bioinformatics* **19**: 2486–2488.
- Li, Y., Korolev, S. and Waksman, G. (1998). Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J* **17**: 7514–7525.
- Li, B.H., Kadura, I., Fu, D.J., and Watson, D.E. (2004) Genotyping with TaqMAMA. *Genomics* **83**: 311–320.

- Liang, P. and Pardee, B.A. (1992). Differential display of eukaryotic mRNA by means of polymerase chain reaction. *Science* **257**: 967–971.
- López-Calleja, I.M., de la Cruz, S., González, I., García, T. and Martín, R. (2005). Market analysis of food products for detection of allergenic walnut (*Juglans regia*) and pecan (*Carya illinoensis*) by real-time PCR. *Food Chem* **177**: 111–119.
- Markham, N.R. and Zuker, M. (2005). DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Research* **33**(Web Server issue): W577–W581.
- Marky, L.A. and Breslauer, K. J. (1982). Calorimetric determination of base-stacking enthalpies in double-helical DNA molecules. *Biopolymers* **21**(11): 2185–2194.
- Marshall, O.J. (2004). PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* **20**: 2471–2472.
- Nielsen, H.B. and Knudsen, S. (2002). Avoiding cross hybridization by choosing nonredundant targets on cDNA arrays. *Bioinformatics* **18**, 321–322.
- Oggioni, M.R., Meacci, F., Carattoli, A., Ciervo, A., Orru, G., Cassone, A. and Pozzi, G. (2002). Protocol for real-time PCR identification of anthrax spores from nasal swabs after broth enrichment. *J Clin Microbiol* **40**(11): 3956–3963.
- Peyret, N., Seneviratne P.A., Allawi, H.T. and SantaLucia, J.Jr. (1999). Nearest-neighbor thermodynamics and NMR of DNA sequences with internal AA, CC, GG, and TT mismatches. *Biochemistry* **38**: 3468–3477.
- Piao, X., Yan, Y., Yan, J. and Guan, Y. (2009). Enhanced recognition of noncomplementary hybridization by single-LNA-modified oligonucleotide probes. *Anal Bioanal Chem* **394**: 1637–1643.
- Pienaar, E., Theron, M., Nelson, M. and Viljoen, H.J. (2006). A quantitative model of error accumulation during PCR amplification. *Comput Biol Chem* **30**: 102–111.
- Qu, W., Zhou, Y., Zhang, Y., Lu, Y., Wang, X., Zhao, D., Yang, Y. and Zhang, C. (2012). MFEprimer-2.0: a fast thermodynamics-based program for checking PCR primer specificity. *Nucleic Acids Res* **40**(Web Server issue): W205–W208.
- Rouchka, E.C., Khalyfa, A. and Cooper, N.G. (2005). MPrime: efficient large scale multiple primer and oligonucleotide design for customized gene microarrays. *BMC Bioinformatics* **6**: 175.
- Roux, K.H. (2009). Optimization and troubleshooting in PCR. *Cold Spring Harb. Protoc* pdb ip66.
- Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**(3): 365–386.
- SantaLucia, J.Jr. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA* **95**: 1460–1465.

- SantaLucia, J.Jr. and Hicks, D. (2004). The Thermodynamics of DNA Structural Motifs. *Annu Rev Biophys Biomol Struct* **33**: 415-440.
- SantaLucia, J.Jr (2007). Physical principles and visual-OMP software for optimal PCR design. *Methods Mol Biol* **402**: 3-34.
- Sarkar, G., Cassady, J., Bottema, C.D. and Sommer S. S. (1990). Characterization of polymerase chain reaction amplification of specific alleles. *Anal Biochem* **186**: 64–68.
- Srivastava, G.P., Guo, J.Y., Shi, H.D. and Xu, D. (2008). PRIMEGENS-v2: genome-wide primer design for analyzing DNA methylation patterns of CpG islands. *Bioinformatics* **24**(17): 1837–1842.
- Stadhouders, R., Pas, S.D., Anber, J., Voermans, J., Mes, T.H.M. and Schutten, M. (2010). The effect of primertemplate mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay. *J Mol Diagn* **12**: 109–117.
- Thweatt, R., Goldstein, S. and Reis, R.J.S. (1990). A universal primer mixture for sequence determination at the 3' ends of cDNAs. *Anal Biochem* **190**: 314–316.
- Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J.A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* **35**(Web Server issue): 71–74.
- Untergasser, A., Cutcutache, L., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res* **40**(15): e115.
- Van Borm, S. and Boomsma, J.J. (2002). Group-specific polymerase chain reaction amplification of SSU rRNA-encoding gene fragments from 12 microbial taxa. *Molecular Ecology Notes* **2**: 356-359.
- Villard, P. and Malausa, T. (2013). SP-Designer: a user-friendly program for designing species-specific primer pairs from DNA sequence alignments. *Mol Ecol Resour* **13**: 755–758.
- Watson, J.D. (1971). The regulation of DNA synthesis in eukaryotes. *Adv Cell Biol* **2**: 1–46.
- Watson, R. (1989). The formation of primer artifacts in polymerase chain reactions. *Amplifications* **1**: 5–6.
- Weckx, S., De Rijk, P., Van Broeckhoven, C. and Del Favero, J. (2005). SNPbox: a modular software package for large-scale primer design. *Bioinformatics* **21**: 385-387.
- Whitney, S.E., Sudhir, A., Nelson, R.M. and Viljoen, H.J. (2004). Principles of rapid polymerase chain reactions: mathematical modeling and experimental verification. *Comput Biol Chem* **28**: 195–209.
- Wiesner, R.J. (1992). Direct quantification of picomolar concentrations of mRNAs by mathematical analysis of a reverse transcription/exponential polymerase chain reaction assay, *Nucleic Acids Res* **20**: 5863–5864.

- Wilcox, T.M., McKelvey, K.S., Young, M.K., Jane, S.F., Lowe, W.H., Whiteley, A.R. and Schwartz, M.K. (2013). Robust detection of rare species using environmental DNA: the importance of primer specificity. *PLOS ONE* **8**(3): e59520.
- Wright, E.S., Yilmaz, L.S., Ram, S., Gasser, J.M., Harrington, G.W. and Noguera, D.R. (2014). Exploiting extension bias in polymerase chain reaction to improve primer specificity in ensembles of nearly identical DNA templates. *Environmental Microbiology* **16**: 1354–1365.
- Wu, J.H., Hong, P.Y. and Liu, W.T. (2009). Quantitative effects of position and type of single mismatch on single base primer extension. *J Microbiol Methods* **77**: 267–275.
- Xu, D., Li, G., Wu, L., Zhou, J. and Xu, Y. (2002). PRIMEGENS: robust and efficient design of gene-specific probes for microarray analysis. *Bioinformatics* **18**:1432-1437.
- Yang, C.H., Cheng, Y.H., Yang, C.H. and Chuang, L.Y. (2012). Mutagenic primer design for mismatch PCR-RFLP SNP genotyping using a genetic algorithm. *IEEE/ACM Trans Comput Biol Bioinform* **9**(3):837–845.
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. and Madden, T. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**(1): 134.
- Zhao, G. and Guan, Y. (2010). Polymerization behavior of Klenow fragment and Taq DNA polymerase in short primer extension reactions. *Acta Biochim Biophys Sin (Shanghai)* **42**(10):722–728.

Kasutatud veebiaadressid

<http://genome.ucsc.edu/cgi-bin/hgPcr?command=start> (aprill 2015)

<http://www.dnasoftware.com/problems-we-solve/thermoblast> (aprill 2015)

Lisad

Lisa 1

Mittepaarduvate nukleotiidide asukohad makadaamiapähkli spetsiifiliste praimerite alusel genereeritud pärisuunalistes praimerites (vastassuunaline praimer on täies pikkuses komplementaarne sihtmärkjärjestusega). Makadaamiapähkli praimeritega identsed positsioonid on tähistatud tärniga (*) ning varieeruvad (mittepaarduvate nukleotiididega) positsioonid tähistatud tähega N (moodustuv mittepaarduv aluspaar on nõrga jõuga, puriin/pürimidiin või pürimidiin/puriin tüüpi). $\Delta\Delta G$ väärtus tähistab praimeri Gibbsi vaba energia muudu erinevust mittepaarduvate nukleotiididega (ΔG_m) ja täies pikkuses sihtmärkjärjestusega komplementaarse praimeri (ΔG_0) seondumise vahel. Makadaamiapähkli täies pikkuse komplementaarse pärisuunalise praimeri $\Delta G_0 = -24,3$. NAS (*No Amplification Score*) väärtus on arvutatud töös kirjeldatud valemi alusel (mittepaarduvate nukleotiidide arvu summa 20, 10, 6 ja 3 nukleotiidi hulgas praimeri 3'-otsast). PCR saagist on hinnatud silma järgi PCR produktide geelelektroforeesipildi triibu (*band*) intensiivsuse alusel (0 – triip ei ole tuvastatav, 1 – triip on oluliselt nõrgem võrreldes positiivse kontrolliga, 2 – triibu intensiivsus tugev, võrreldav positiivse kontrolliga).

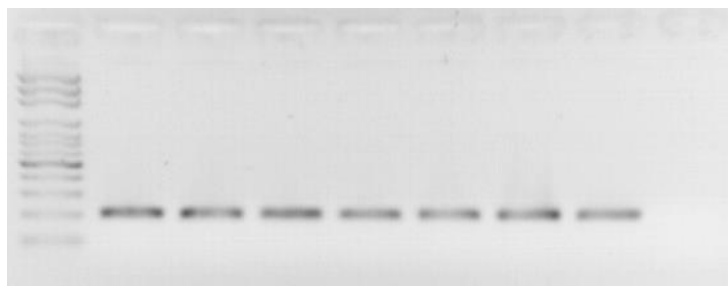
Praimeri ID	Praimeri järjestus 5'→3'	ΔG_m	$\Delta\Delta G$	NAS	PCR saagis
Mc21A_F	****N*N*****N*N*N**NNNN*	-7,5	16,8	21	0
Mc20A_F	****N*N*****N*N*N**NNNN*	-7,3	17	20	0
Mc19A_F	****N*****NN**NN*N**N	-10,9	13,4	19	0
Mc19B_F	****N*N*NN**N*N*N**NN**N	-6,9	17,4	19	0
Mc18A_F	*****NNNNN	-20,3	4	18	0
Mc18B_F	*****NN**NN*N**N	-14,5	9,8	18	0
Mc18C_F	****N*N**NNN**NNN*N*****N	-6,2	18,1	18	0
Mc17A_F	*****N***N**NNN	-16,5	7,8	17	0
Mc17B_F	****N*****NN*N**NN**N*	-10,9	13,4	17	0
Mc17C_F	****NN**N*N**N**N**N*N	-6,7	17,6	17	0
Mc16A_F	*****N***NN*N*N	-16,5	7,8	16	0
Mc16B_F	****N*****N**NN**N*N	-12,5	11,8	16	0
Mc16C_F	****N**NN*****NNN*N**N**	-5,8	18,5	16	0
Mc15A_F	*****NNNNN	-21,5	2,8	15	0
Mc15B_F	*****N**NN**N*N	-16,1	8,2	15	0
Mc15C_F	****N***N*****NNN*N**N**	-6,7	17,6	15	0
Mc14A_F	*****NN**NN	-18,4	5,9	14	0
Mc14B_F	****N*****N***N**NN*	-13,0	11,3	14	0
Mc13A_F	*****N***N***NN	-16,5	7,8	13	0
Mc13B_F	*****N***N***NN*	-16,5	7,8	13	0
Mc12A_F	*****NNN	-21,5	2,8	12	1
Mc12B_F	*****N***NN***N	-16,5	7,8	12	1
Mc12C_F	****N*****NN**N***N*	-10,9	13,4	12	1
Mc11A_F	*****N**N*N	-18,4	5,9	11	1
Mc11B_F	****N***N***N***N***N	-9,0	15,3	11	1
Mc11C_F	****N***N***N***N***N**	-8,6	15,7	11	1
Mc10A_F	*****N*****N*N	-19,6	4,7	10	1
Mc10B_F	*****N*****N***N***N	-12,8	11,5	10	0
Mc10C_F	****N*****N***N***N**	-13,0	11,3	10	1
Mc10D_F	****N***N***NN*N*N*****	-5,3	19	10	0
Mc10E_F	NNNNNNNNNNNNNNNN*****	-9,1	15,2	10	1

Mc9A_F	*****N***N***N	-17,0	7,3	9	1
Mc9B_F	*****N*****N*N	-17,8	6,5	9	0
Mc9C_F	****N*****NN*	-17,9	6,4	9	0
Mc9D_F	****N*****N*N*****N	-12,3	12	9	1
Mc9E_F	****N*****N**N***N*	-14,5	9,8	9	1
Mc9F_F	NNNNNNNNNNNN*****	-9,4	14,9	9	1
Mc8A_F	*****N*N	-21,5	2,8	8	1
Mc8B_F	*****NN*	-21,5	2,8	8	1
Mc8C_F	****N*****NN*****N	-14,4	9,9	8	1
Mc8D_F	****NN***N*****N*N*****	-6,7	17,6	8	0
Mc8E_F	NNNNNNNNNNNN*****	-11,1	13,2	8	1
Mc7A_F	*****N***N	-18,8	5,5	7	1
Mc7B_F	*****N***N*	-18,4	5,9	7	1
Mc7C_F	*****N*N**	-18,4	5,9	7	1
Mc7D_F	****N*****N*****N	-17,8	6,5	7	1
Mc7E_F	NNNNNNNNNNNN*****	-11,2	13,1	7	1
Mc6A_F	*****N*N***	-18,4	5,9	6	1
Mc6B_F	*****NN*****	-18,9	5,4	6	1
Mc6C_F	*****N*****N	-21,3	3	6	1
Mc6D_F	*****N*****N*	-20,4	3,9	6	1
Mc6E_F	*****N*****N**	-19,6	4,7	6	1
Mc6F_F	****N*****N*****N	-16,0	8,3	6	1
Mc6G_F	NNNNNNNNNNNN*****	-13,3	11	6	1
Mc5A_F	*****N*****N***	-20,5	3,8	5	2
Mc5B_F	*****N*****N*****	-19,0	5,3	5	2
Mc5C_F	*****N***N*****	-18,1	6,2	5	2
Mc5D_F	*N*****N	-19,8	4,5	5	1
Mc5E_F	*N*****N*	-18,9	5,4	5	1
Mc5F_F	*N*****N**	-18,1	6,2	5	2
Mc5G_F	NNNNNNNNNN*****	-13,4	10,9	5	2
Mc4A_F	*****N	-23,2	1,1	4	2
Mc4B_F	*****N*	-22,2	2,1	4	2
Mc4C_F	*****N**	-21,5	2,8	4	2
Mc4D_F	*N*****N***	-19,0	5,3	4	2
Mc4E_F	*N*****N*****	-17,5	6,8	4	2
Mc4F_F	*N*****N*****	-16,6	7,7	4	2
Mc4G_F	NNNNNNNN*****	-14,0	10,3	4	2
Mc3A_F	*****N***	-22,3	2	3	2
Mc3B_F	*****N*****	-20,9	3,4	3	2
Mc3C_F	*****N*****	-19,9	4,4	3	2
Mc3D_F	*N*****N*****	-19,1	5,2	3	2
Mc3E_F	NNNNNNNN*****	-14,7	9,6	3	2
Mc2A_F	NNNNNN*****	-17,1	7,2	2	2
Mc1A_F	NNNNN*****	-17,2	7,1	1	2
Mc0A_F	NNNN*****	-17,7	6,6	0	2
Mc0_F	*****	-24,3	0	0	2

Lisa 2

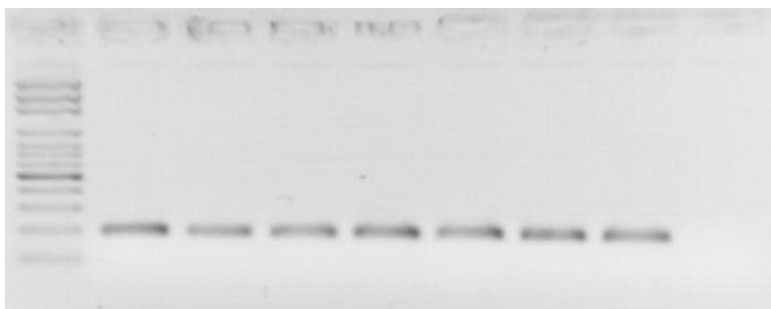
Amplifikatsioonireaktsioonide tulemused praimeripaaridega, mille pärisuunalises praimeris on üks erinevas positsioonis (1.-6.) paiknev mittepaarduv nukleotiid ning vastassuunaline praimer on täies pikkuses komplementaarne sihtmärkjärjestusega. a) pärisuunalise praimer mittepaarduv nukleotiid on 1. positsioonis (praimer 3'-otsast), moodustades erinevat tüüpi mittepaarduvaid aluspaare (GT, TT, CT). Positiivses kontrollis on 1. positsioonis nukleotiidide vahel Watson-Crick paardumine AT, mille puhul praimer 3'-otsas on A-nukleotiid ning sihtmärkjärjestuse vastavas positsioonis on T-nukleotiid. b) - f) pärisuunalise praimer erinev mittepaarduv nukleotiid paikneb vastavalt 2. - 6. positsioonis. g) amplifikatsioon üksiku erinevas positsioonis paikneva nõrgema mõjuga (puriin/pürimidiin või pürimidiin/puriin tüüpi) mittepaardunud aluspaari esinemisel praimer ja sihtmärkjärjestuse vahel. h) amplifikatsioon üksiku erinevas positsioonis paikneva tugevama mõjuga (puriin/puriin või pürimidiin/pürimidiin tüüpi) mittepaarduva aluspaari esinemisel praimer ja sihtmärkjärjestuse vahel. Lisatud on negatiivne ja positiivne kontroll (Mc0).

a)



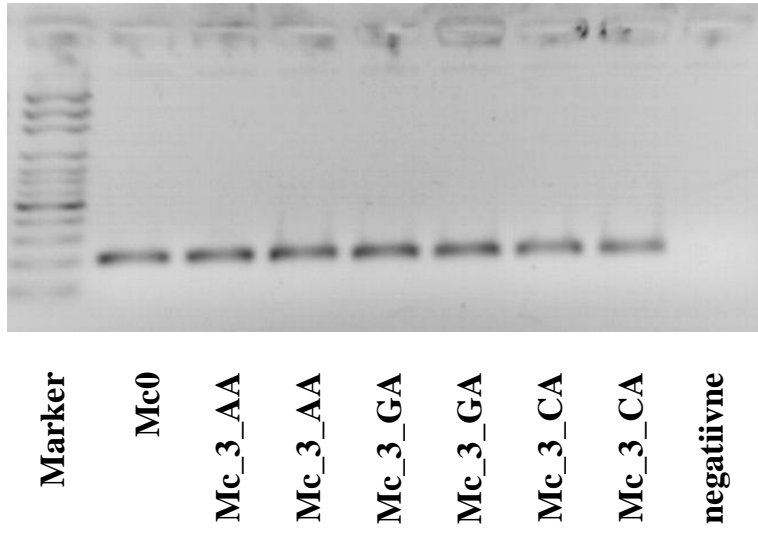
Marker
Mc0
Mc_1_GT
Mc_1_CT
Mc_1_TT
Mc_1_TT
Mc_1_TT
Mc_1_CT
Mc_1_CT
Negatiivne

b)

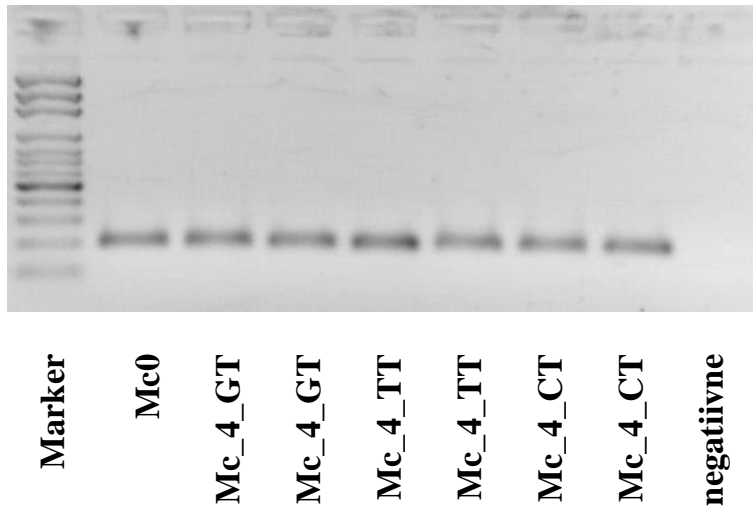


Marker
Mc0
Mc_2_TG
Mc_2_TG
Mc_2_GG
Mc_2_GG
Mc_2_AG
Mc_2_AG
negatiivne

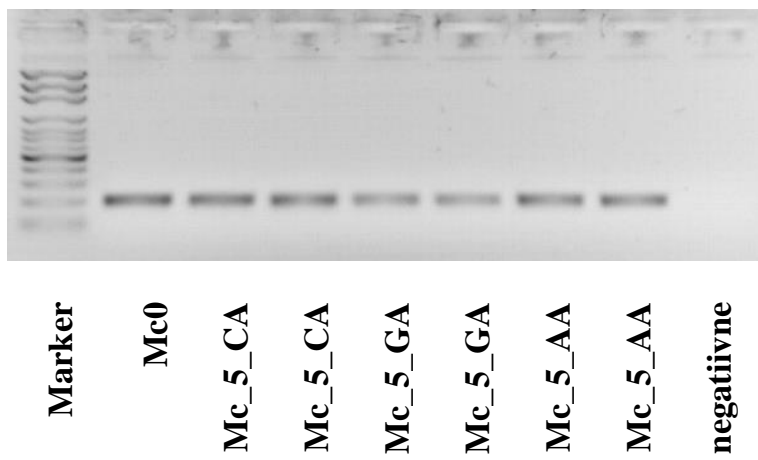
c)



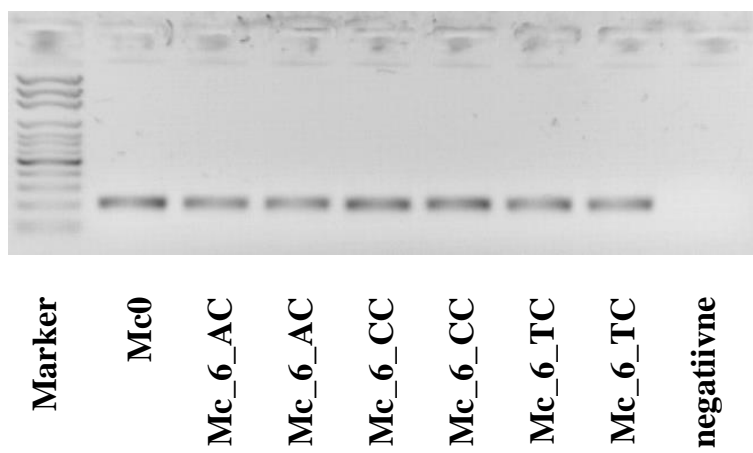
d)



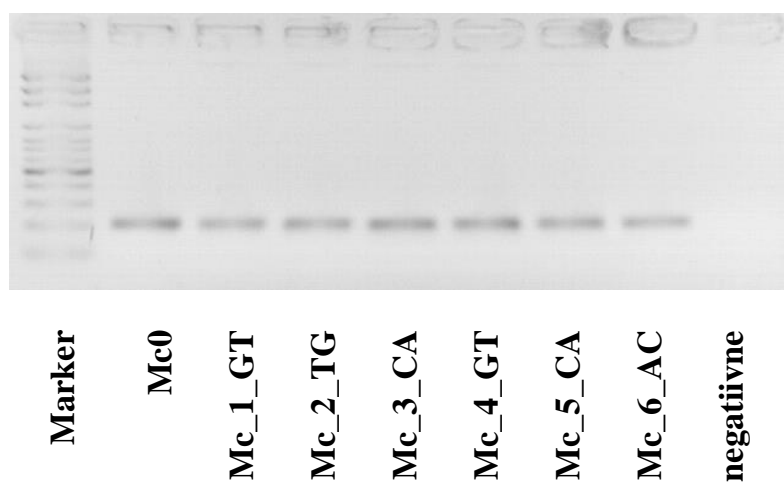
e)



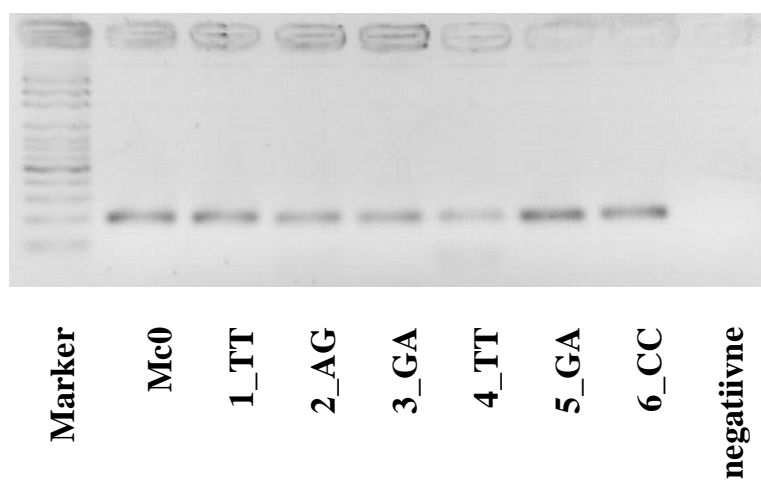
f)



g)



h)



Lisa 3

Amplifikatsiooni tulemus praimeripaariga, mille pärisuunalises praimeris on kaks mittepaarduvat nukleotiidi ja vastassuunaline priamer on täies pikkuses komplementaarne sihtmärkjärjestusega. a) 1. positsioonis (praimer 3'-otsas) paiknevale tugeva (T) või nõrga mõjuga (N) mittepaardunud aluspaarile lisaks on teine nõrga mõjuga mittepaarduv aluspaar 3., 6., 10. või 23. positsioonis praimer 3'-otsast. b) – f) vastavalt 2. - 6. või 10. positsioonis paiknevale tugeva (T) või nõrga mõjuga (N) mittepaardunud aluspaarile lisaks on teine nõrga mõjuga mittepaarduv aluspaar 3., 6., 10. või 23. Positsioonis. Lisatud on negatiivne ja positiivne kontroll (Mc0).

a)



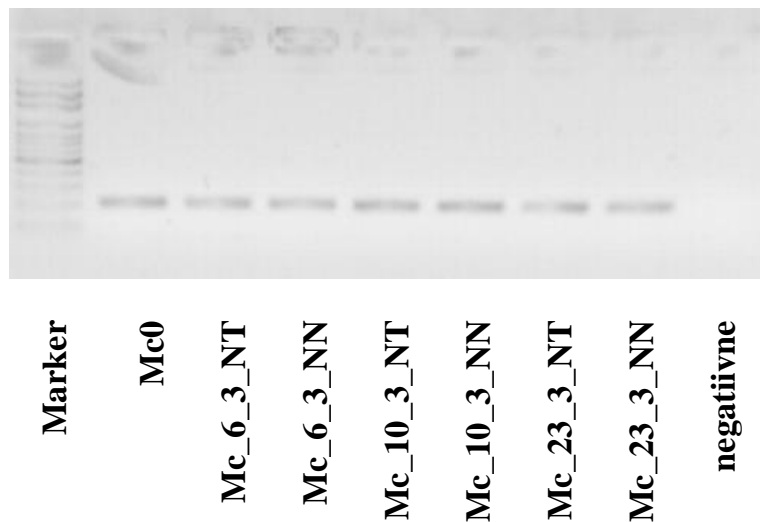
Marker	Mc0	Mc_3_1_NT	Mc_3_1_NN	Mc_6_1_NT	Mc_6_1_NN	Mc_10_1_NT	Mc_10_1_NN	Mc_23_1_NT	Mc_23_1_NN	negatiivne
---------------	------------	------------------	------------------	------------------	------------------	-------------------	-------------------	-------------------	-------------------	-------------------

b)

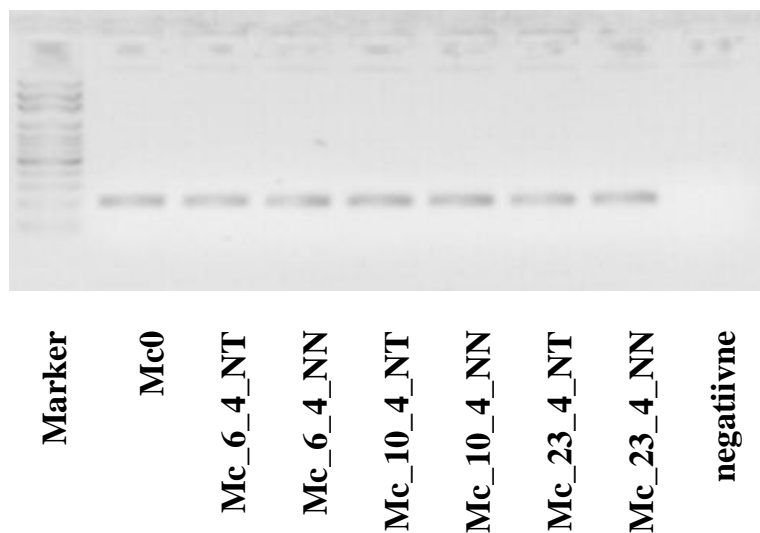


Marker	Mc0	Mc_3_2_NT	Mc_3_2_NN	Mc_6_2_NT	Mc_6_2_NN	Mc_10_2_NT	Mc_10_2_NN	Mc_23_2_NT	Mc_23_2_NN	negatiivne
---------------	------------	------------------	------------------	------------------	------------------	-------------------	-------------------	-------------------	-------------------	-------------------

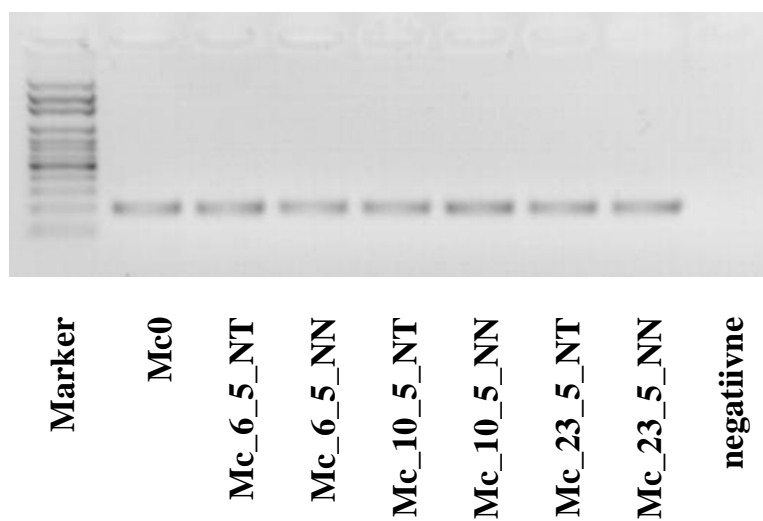
c)



d)



e)



f)



Marker

Mc0

Mc_10_6_NT

Mc_10_6_NN

Mc_23_6_NT

Mc_23_6_NN

Mc_23_10_NT

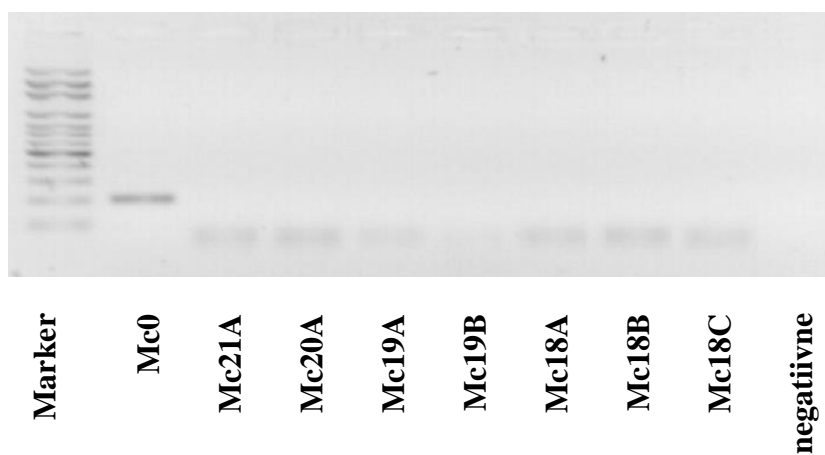
Mc_23_10_NN

negativne

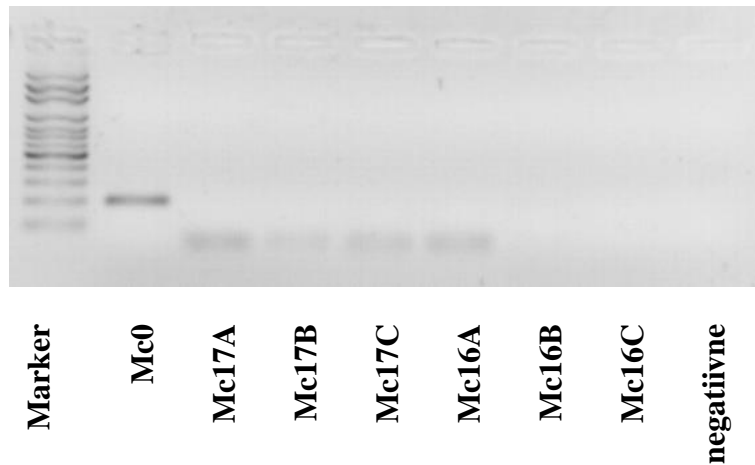
Lisa 4

Amplifikatsiooni tulemus erineva NA skoori väärtusega praimeripaaride puhul. a) praimeripaaride NAS väärtus on 18 - 21, b) praimeripaaride NAS väärtus on 16 - 17, c) praimeripaaride NAS väärtus on 13 - 15, d) praimeripaaride NAS skoori väärtus on 11 - 12, e) praimeripaaride NAS skoori väärtus on 9 - 10, f) praimeripaaride NAS skoori väärtus on 8 - 9, g) praimeripaaride NAS skoori väärtus on 7 - 8, h) praimeripaaride NAS skoori väärtus on 6, i) praimeripaaride NAS skoori väärtus on 5, j) praimeripaaride NAS skoori väärtus on 4, k) praimeripaaride NAS skoori väärtus on 0 - 3. Lisatud on positiivne kontroll (Mc0, mille puhul NAS väärtus on 0) ning negatiivne kontroll.

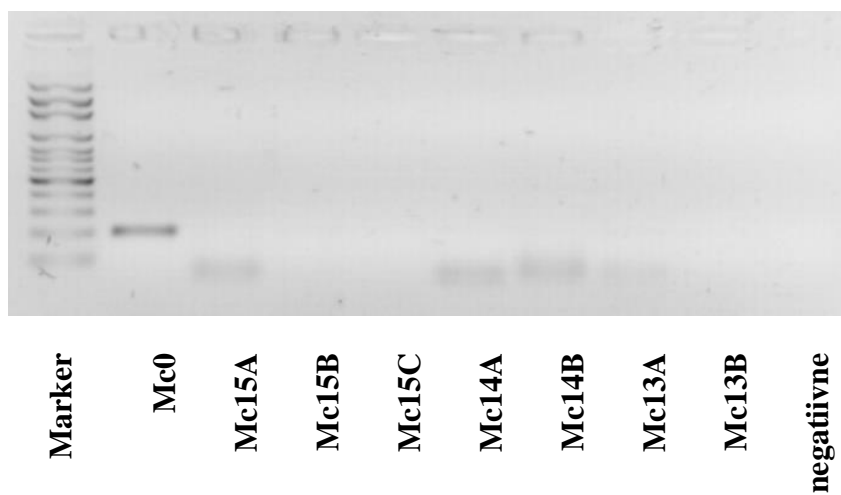
a)



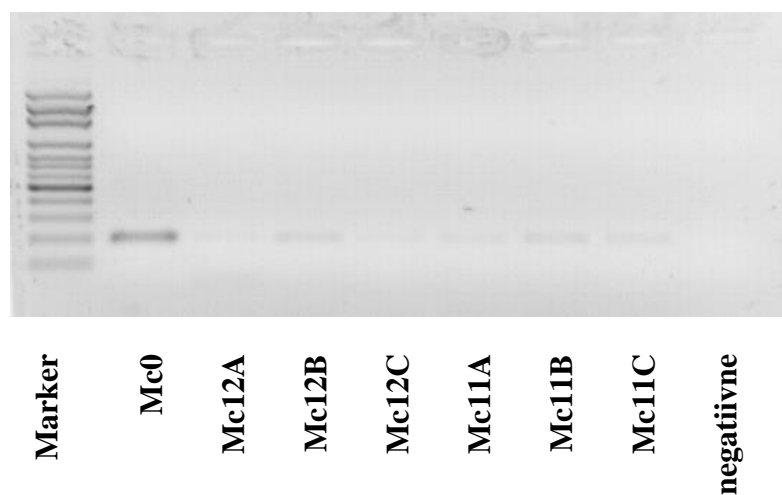
b)



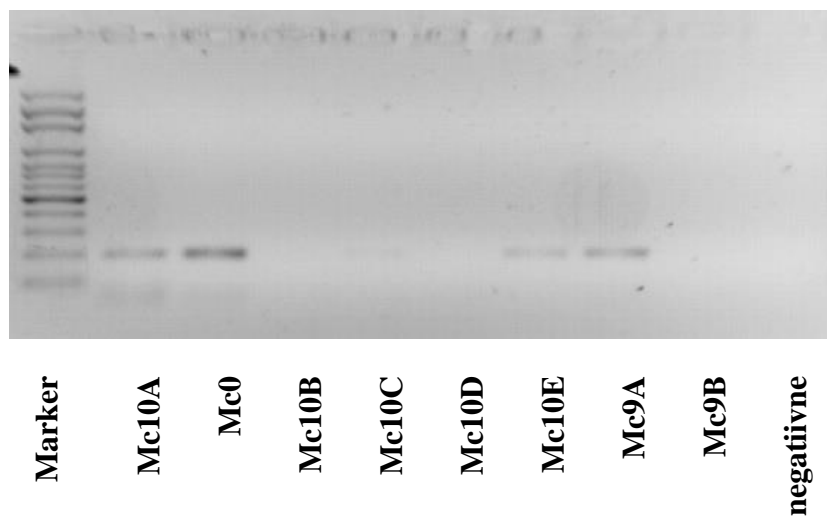
c)



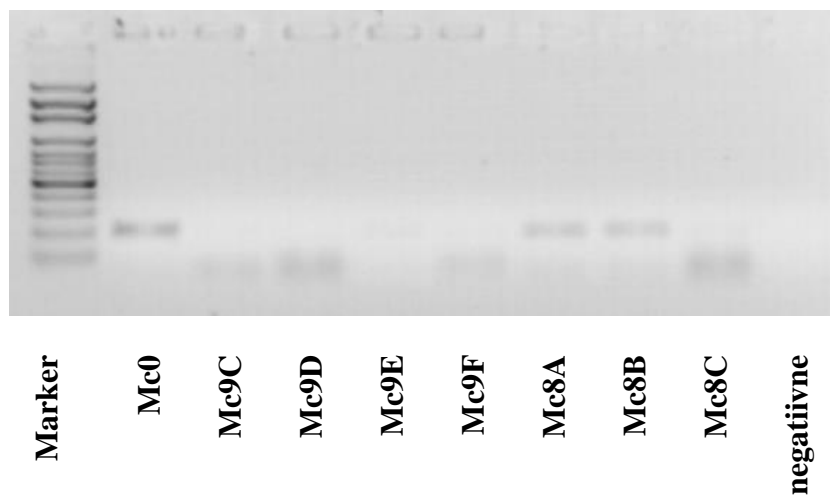
d)



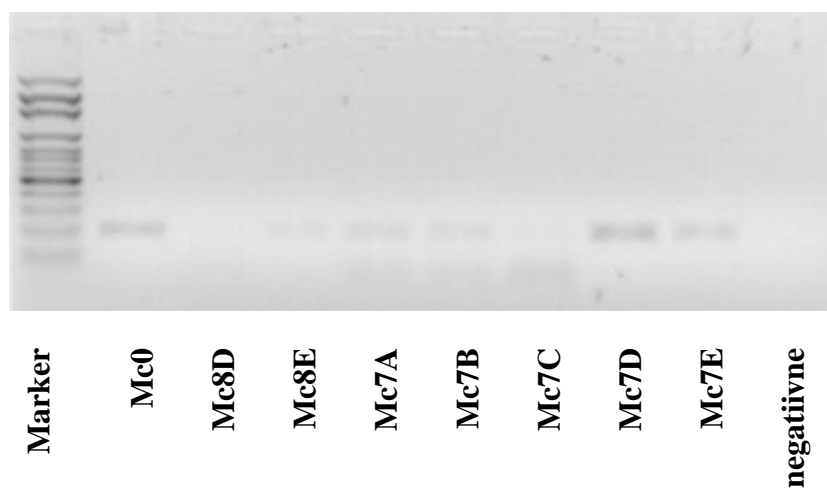
e)



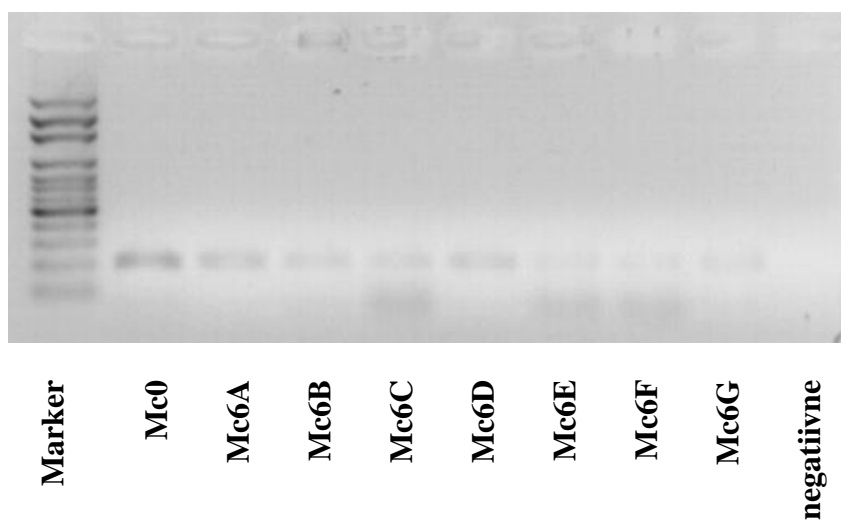
f)



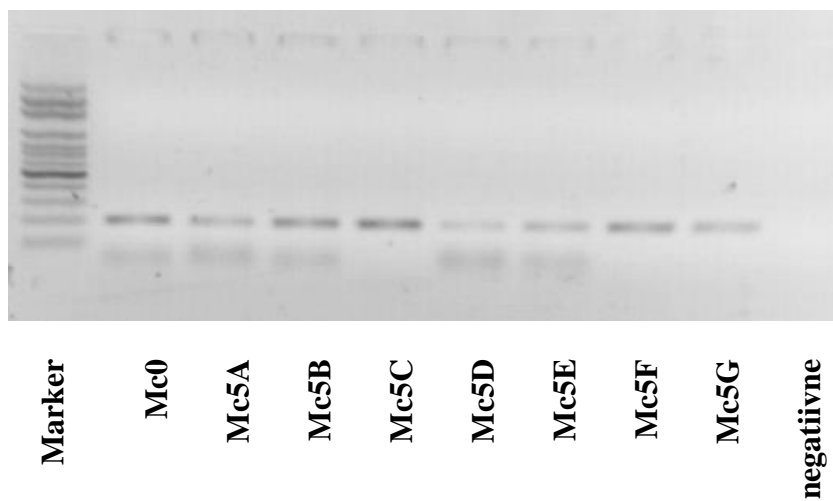
g)



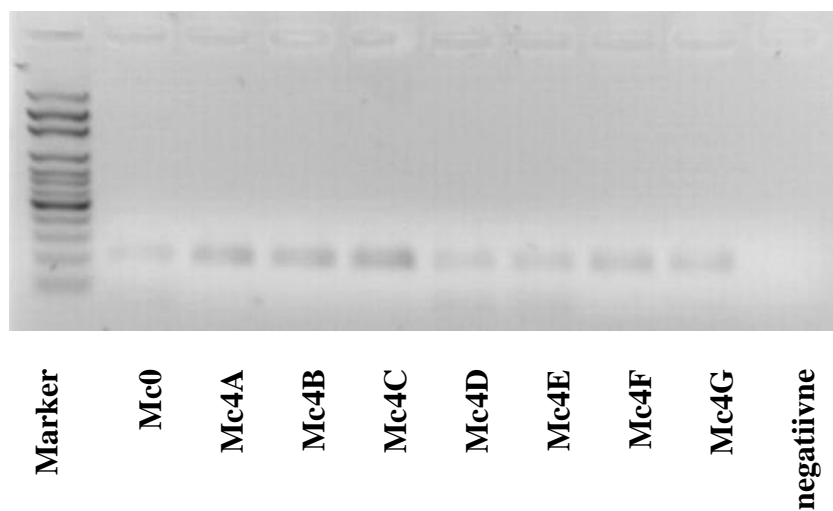
h)



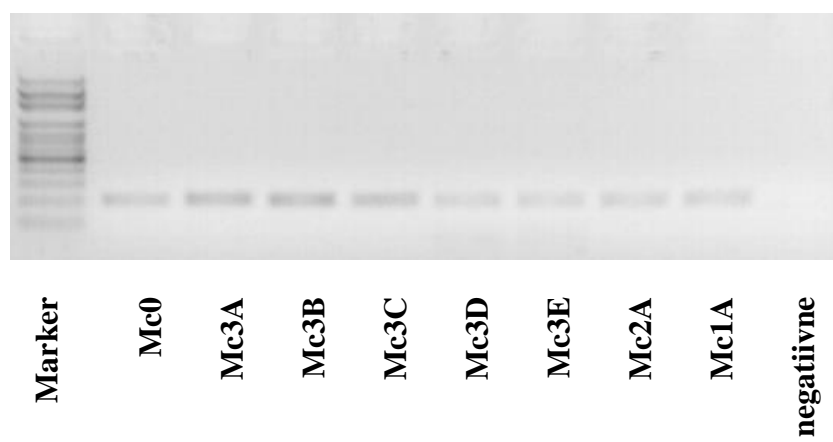
i)



j)



k)



Lisa 5

Pekani-hikkoripuu spetsiifiliste praimerite disainil kasutatavad programmi Praimer3 parameetri väärtused.

PRIMER_PRODUCT_SIZE_RANGE=70-300

PRIMER_OPT_SIZE=22

PRIMER_MIN_SIZE=18

PRIMER_MAX_SIZE=27

PRIMER_OPT_TM=59

PRIMER_MIN_TM=57

PRIMER_MAX_TM=60

PRIMER_PAIR_MAX_DIFF_TM=4

PRIMER_OPT_GC_PERCENT=35

PRIMER_MIN_GC=10

PRIMER_MAX_GC=90

PRIMER_SALT_DIVALENT=2.5

PRIMER_DNTP_CONC=0.25

PRIMER_TM_FORMULA=1

PRIMER_SALT_CORRECTIONS=1

P3_FILE_FLAG=0

PRIMER_EXPLAIN_FLAG=1

PRIMER_MAX_POLY_X=4

PRIMER_NUM_RETURN=800

PRIMER_MAX_HAIRPIN_TH=40

PRIMER_MAX_NS_ACCEPTED=0

PRIMER_THERMODYNAMIC_ALIGNMENT=1

PRIMER_LOWERCASE_MASKING=1

PRIMER_MIN_THREE_PRIME_DISTANCE=0

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Kairi Raime (sünnikuupäev: 08.06.1982)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Taksoni-spetsiifiliste praimerite disaini meetodika arendus ning selle rakendamine allergeensete taimeliikide tuvastamisel“,

mille juhendajad on prof. Maido Remm ja Triinu Kõressaar, PhD

- 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu alates 31.12.2017 kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 26.05.2015